مقدمة في الاقتصاد القياسي

تأليف الأستاذ الدكتور

محمد صالح تركي القريشي

الطبعة الأولى

2004



الإهداء

* الى أساتذة تعلمت منهم
* الى طلابي
* الى اسرتي زوجتي وبناتي
وابني علاء

المحتويات

الموضوع	الصفحة
المقدمة	9
الفصل الأول / تعريف علم الإقتصاد القياسي	13
الإقتصاد القياسي بوصفه موضوعاً مستقلاً	14
طبيعة علم الإقتصاد القياسي	16
أغراض الإقتصاد القياسي	17
ما هو النموذج ؟	23
أنواع النماذج	26
أنواع العلاقات الإقتصادية	31
منهجية البحث العلمي في الإقتصاد القياسي	32
أنواع الإقتصاد القياسي	**37**
الفصل الثاني / نظرية الإرتباط	39
مقدمة	41
قياس الإرتباط الخطي	44
معامل إرتباط الرتب	54
معاملات الإرتباط الجزئي	**56**
حدود نظرية الإرتباط الخطي	58
الفصل الثالث / طبيعة تحليل الإنحدار	63
الأصل التاريخي لمصطلح الإنحدار	**65**
التفسير الحديث للإنحدار	65
الإعتماد الإحصائي مقابل الإعتماد الدالي	69
الفصل الرابع / تحليل الإنحدار البسيط	71
مقدمة	73
مفهوم دالة إنحدار المجتمع الإحصائي	79
التحديد الإحتمالي لدالة إنحدار المجتمع الإحصائي	81
دالة إنحدار العينة	83
الفصل الخامس / نموذج الإنحدار الخطي البسيط وطريقة المربعات الصغرى الإعتيادية	91

المحتويات

الموضوع	الصفحة
مقدمة	**93**
الموضوع	**الصفحة**
نموذج الإنحدار الخطي البسيط	94
إفتراضات نموذج الإنحدار الخطي الإحتمالي	**104**
معيار المربعات الصغرى والمعادلات الإعتيادية للمربعات الصغرى الإعتيادية	112
إشتقاق المعادلات الطبيعية أو الإعتيادية	120
تقدير الدالة التي يكون تقاطعها مع المحور العمودي صفراً	127
تقدير المرونات من خط الإنحدار المقدر	**128**
الفصل السادس / الإنحدار المتعدد	131
نموذج ذو متغيرين توضيحيين	133
نموذج العلاقات غير الخطية في الإنحدار	143
تحويل الدوال غير الخطية التي تحتوي مرونات ثابتة الى دوال خطية	145
الفصل السابع / الإختبارات الإحصائية لمعنوية المعلمات المقدرة بطريقة المربعات الصغرى الإعتيادية	147
مقدمة	149
إختبار جودة توفيق خط الإنحدار بإستخدام مربع معامل الإرتباط (r^2)	150
إختبار الخطأ المعياري للمعلمات المقدرة ($\hat{b}_1 \quad \hat{b}_0$) بطريقة المربعات الصغرى الإعتيادية	154
إختبار (Z) للمعلمات المقدرة بطريقة المربعات الصغرى	157
إختبار (t) للمعلمات المقدرة بطريقة المربعات الصغرى	**163**
حدود الثقة للمعلمات($\hat{b}_1 \quad \hat{b}_0$)	168
الفصل الثامن / تحليل التباين	172
مقدمة	175
طريقة تحليل التباين بوصفها طريقة إحصائية	177
إختبار الفرق بين المتوسطات الحسابية	**179**
إختبار (F) لمعنوية خط الإنحدار	193
الفصل التاسع مشكلات الإنحدار / مشكلة الإرتباط الخطي المتعدد	**197**

المحتويات

مقدمة	199
أسباب ظهور أو وجود مشكلة الإرتباط الخطي المتعدد	200
الموضوع	**الصفحة**
ما يترتب على وجود الإرتباط الخطي المتعدد	202
إختبارات للكشف عن وجود الإرتباط الخططي المتعدد	**208**
بعض الحلول لمشكلة الإرتباط الخطي المتعدد	224
الفصل العاشر / مشكلات الإنحدار / مشكلة الإرتباط الذاتي	229
مقدمة	231
طبيعة مشكلة الإرتباط الذاتي	231
النتائج المترتبة على وجود مشكلة الإرتباط الذاتي	238
إختبارات الإرتباط الذاتي	239
بعض النواقص في إختبار دربن ــ واطسن	244
الحلول لمشكلة الإرتباط الذاتي	244
الفصل الحادي عشر / مشكلات الإنحدار / مشكلة عدم ثبات التباين	247
مقدمة	249
التفسير والعرض بصيغة شكل الإنتشار لثبات وعدم ثبات التباين	249
إمكانية وجود إفتراض تجانس التباين	251
النتائج المترتبة على وجود مشكلة عدم تجانس أو عــدم ثبــات تبــاين المتغيــر العشوائي (U)	253
إختبارات ثبات أو عدم تجانس التباين	254
الحلول لمشكلة عدم ثبات التباين	258
الفصل الثاني عشر / إسلوب المصفوفات في تحليل الإنحدار	261
مقدمة	263
المصفوفات	263
أمثلة إنحدار	**268**
تساوي المصفوفات	**269**
جمع وطرح المصفوفات	**271**

المحتويات

ضرب المصفوفات	273
أمثلة في إستخدام المصفوفات في تحليل الإنحدار	277
أنواع خاصة من المصفوفات	279
الموضوع	**الصفحة**
مثال على تحليل الإنحدار المتعدد بمتغيرين توضيحيين بإستخدام المصفوفات	285
بعض النظريات المهمة للمصفوفات	290
المتجهات والمصفوفات العشوائية	**290**
مصفوفة التباين والتباين المشترك للمتجه العشوائي	**291**
نموذج الإنحدار الخطي البسيط بصيغ المصفوفات	**292**
تقدير معلمات الإنحدار بطريقة المربعات الصغرى	294
تحليل التباين	296
الفصل الثالث عشر / نماذج المعادلات الآنية	301
مقدمة	303
طبيعة نماذج المعادلات الآنية	304
مثال على نماذج المعادلات الآنية	306
حل تحيز المعادلات الآنية	309
الفصل الرابع عشر / التشخيص الإحصائي	323
مقدمة	325
مشكلة التشخيص الإحصائي	332
الفصل الخامس عشر / تطبيقات وتمارين محلولة	347
بعض التمارين المحلولة	349
بعض التطبيقات الإقتصادية القياسية	356
المراجع	363

المقدمة

الحمد لله الذي مكنني من إتمام هذا العمل ، لقد كان مشروع تأليف كتاب في موضوع الإقتصاد القياسي أحد أهم أهدافي في عملي العلمي إذ بدأت أفكر في هذا المشروع منذ اليوم الأول الذي بدأت به تدريس هذا الموضوع عندما كنت أستاذاً في جامعة الموصل عام 1985 ولغاية عام 1997 ، وثم في الجامعة المستنصرية منذ عام 1997 ولغاية عام 2002 ، ولكن الظروف كانت غير ملائمة لإنجاز هذا المشروع المهم . إن السبب الرئيس في إهتمامي بهذا المشروع هو الحاجة الماسة الى مصدر باللغة العربية يساعد الطلبة والباحثين في فهم هذا الموضوع وتطبيقاتة بوصفه مادة علمية تتمضن منهجية نظرية وتجريبية مفيدة جداً في مجالات البحث العلمي في الإقتصاد .

إن هذا الكتاب ليس فقط كتاباً آخر في الإقتصاد القياسي ، إنه يغطي الموضوعات الأساسية في الموضوع ، وبخاصة تحليل الإنحدار البسيط والمتعدد وإستخدام المصفوفات في تحليل الإنحدار والتباين وعرضها بصيغة مناسبة ، وطرق المعادلات الآنية في تحليل الإنحدار ، ومن ناحية أخرى فقد تم إعطاء أهمية كبيرة لمشكلات الإنحدار وشرحها بقدر من التفصيل لأنها تكمل عملية دراسة تحليل الإنحدار .

يستهدف هذا الكتاب جمهوراً من طلبة الدراسات الجامعية الأولية والعليا في الإقتصاد والعلوم المالية والمصرفية فضلاً عن الباحثين في الشؤون والظواهر الإقتصادية ، إذ أصبح الإقتصاد القياسي من الموضوعات المهمة والشائعة بين الموضوعات الدراسية في الجامعات في العالم ، وذلك لإن هذا الموضوع يعد من أدوات البحث العلمي الأساسية ، والذي يربط بين نظرية الإقتصاد القياسي والجوانب التطبيقية على مستوى الواقع وهذا ما يجعل هذا الكتاب عاماً وملائماً ومهماً .

إن هدفي الرئيس من كتابة هذا الكتاب هو أن يكون موضوع الإقتصاد القياسي مقرباً وممكن الفهم من قبل طلبة كليات الإقتصاد والعلوم الإدارية والباحثين ، لقد كتبت موضوعات هذا الكتاب على هدي فكرة التأكيد على الربط بين الأسس النظرية لعلم الإقتصاد القياسي وأهمها النظرية الإقتصادية ، والأدوات الرياضية والإحصائية ، وهذا الأسلوب هو الذي يحفز القاريء أو الطالب أو الباحث على التفكير بعمق أكثر لتطبيق تلك الأفكار والأدوات ، بمعنى إن هذا الأسلوب يعبر عن محاولة لدراسة علم الإقتصاد والظواهر الإقتصادية بين البعد النظري والبعد القياسي (التجريبي) .

ثمة إعتماد في هذا الكتاب على موضوعات تشكل الركيزة الأساسية في الإقتصاد القياسي ، إذ إن هذا الكتاب يوفر إلماماً جيداً بالمفاهيم والأسس النظرية لهذا الموضوع المهم ، وقد إنصرف التأكيد في هذا الكتاب على توفير التوضيح الملائم لمباديء علم الإقتصاد القياسي مما يؤدي الى جعل الطالب يمتلك القدرة على إستعمال النماذج الإقتصادية القياسية في دراسته وتفسير نتائجها ، وبخاصة في مجال إعداد البحوث العلمية ، لقد سعيت في جهدي هذا بقدر المستطاع أن أسلك طريقاً ملائماً وأكثر سلاسة وعملية في عرض المادة العلمية لهذا الموضوع ، جاعلاً هذا الجهد سبيلاً في توصيل المعرفة العلمية في هذا الموضوع وتسهيل عملية الفهم والإستيعاب لمضامينه ومفرداته من قبل الطالب أو الباحث .

أتمنى أن أكون قد وفقت في تقديم جهداً متواضعاً ومفيداً خدمة للتقدم العلمي والمسيرة العلمية في وطننا العربي .

والله ولي التوفيق

الفصل الأول

تعريف علم الإقتصاد القياسي

تعريف علم الإقتصاد القياسي

1-1 تعريف علم الإقتصاد القياسي .

تعرف كلمة القياس الإقتصادي (Econometrics) بأنها عملية القيـاس الإقتصادي (Economic Measurment) على الرغم من أن القياس يشـكل جزءاً مهماً من الإقتصاد القياسي فإن مدى أو مجال الإقتصاد القياسي أوسـع كثيراً من مجرد عملية القياس ، كما يبدو ذلك من النقـاط الأساسـية المتعلقـة بتعريفات الإقتصاد القياسي :

1. إن الإقتصاد القياسي بوصفه نتيجة للنظر الى دور علـم الإقتصـاد يتألف من تطبيقات الإحصاء الرياضي (Mathematical Statistics) علـى البيانات الإقتصادية (Economic Data) لإعطاء الدعم التجريبي للنماذج التي بنيت في الإقتصاد الرياضي (Mathematical Economics) والحصول على نتائج رقمية (Numerical Results).
2. يمكن أن يعرف الإقتصاد القياسي بوصفه التحليل الكمـي لظـاهرة إقتصادية فعلية (Actual Economic Phenomena) ، تأسيساً على النظرية الإقتصادية والمشاهدة وإستخدام طرقاً إحصائية للإستنتاج .
3. ويمكن أن يعـرف الإقتصـاد القياسـي بأنـه العلـم الإجتمـاعي (Social Science) الـذي يـتم فيـه تطبيـق أدوات النظريـة الإقتصـادية والرياضيات والإستدلال الإحصائي على تحليل الظاهرة الإقتصادية .
4. يهـتم علـم الإقتصـاد القياسـي بـالتقرير التجريبـي (Emprircal Determination) للقوانين الإقتصادية .

وهكذا فإن علم الإقتصاد القياسي يعرف بأنه العلم الذي يدرس الغلاقات الإقتصادية بأسلوب كمي مستخدماً النظرية الإقتصادية والأسـلوب الإحصـائي والحقائق المعبر عنها بإحصاءات منقحة (Refined Data) ويعد علم الإقتصاد

القياسي أحد فروع علم الإقتصاد والتي تستخدم الأدوات الإحصائية والرياضـية للحصول على قيم رقمية لمعلمات (Parameters) المتغيرات التي تعبـر عـن العلاقات الإقتصادية .

2-1 الإقتصاد القياسي بوصفه موضوعاً مستقلاً .

يعد علم الإقتصاد القياسي صياغة دقيقة من عناصـر مثـل النظريـة الإقتصـادية (Economic Theory) ، والإقتصـاد الرياضـي والإحصـاء الإقتصـادي (Economic Statistics) ، والإحصـاء الرياضـي (Mathematical Statistics) بوصفه خليط يكـون شـيئاً جديـداً ، ولكـن الإقتصاد القياسي موضوعاً يستحق أن يدرس مستقلاً بذاته للأسباب الآتية :

السبب الأول : إن النظرية الإقتصادية تضع عبارات وفرضيات غالباً ما تكون نوعية في طبيعتها ، مثلاً إن النظرية الإقتصادية الجزئية تقول إذا بقيـت الأشياء الأخرى ثابتة فإن إنخفاض سعر سلعة معينة من المتوقع أن يؤدي الـى زيادة الكمية المطلوبة من تلك السلعة ، وهكذا فإن النظرية الإقتصـادية تتوقـع علاقة سالبة أو علاقة عكسية بين السعر والكمية المطلوبة من السلعة ، ولكـن النظرية نفسها لا تعطي أي قياس رقمي بين المتغيرين ، بمعنى أنها لاتخبرنـا بالكمية التي تزداد بها الكمية المطلوبة نتيجة لتغيير معين في سعر السلعة ، إنها مهمة المختص بالإقتصاد القياسي (The Econometrician) أن يوفر تقديرات رقمية ، بعبارة أخرى تقول أن الإقتصاد القياسي هو الـذي يعطـي المحتـوى التجريبي (Empirical Content) لمعظم النظرية الإقتصادية .

السبب الثاني : إن الإهتمام الرئيس للإقتصاد الرياضي هو التعبير عـن النظرية الإقتصادية بشكل رياضي أي بصيغة معادلات (Equations) رياضية دون إعتبار لقابلية النظرية على القياس (Measurablity) أو على التجريبيـة (Empirical) بينما نجد إن الإقتصاد القياسي يهتم بشـكل رئـيس بالإثبـات

التجريبي (Empirical Verification) للنظرية الإقتصـــادية ، إن المخـــتص بالإقتصاد القياسي غالباً ما يستعمل معادلات رياضية مقترحة مـــن الإقتصـــاد الرياضي ، ولكنه يضع تلك المعادلات بشكل تسهل فيه عملية الإختبار التجريبي (Empirical Testing) وعملية التحويل للمعادلات الرياضية الـــى معـــادلات إقتصاد قياسي تتطلب الكثير من التأصيل والمهارة العلمية .

السـبب الثالـث : إن الإهتمـــام الـــرئيس للإحصـــاء الإقتصـــادي (Economic Statistics) هو في جمع وتمثيل وعرض البيانات الإقتصـــادية بشكل جداول وأشكال بيانية ، وهذا هو عمل المختص بالإحصاء الإقتصـــادي ، فهو المسؤول أولاً عن جمع بيانات عن الناتج القومي الإجمالي (GNP) ، وعن الإستخدام والبطالة والأسعار ، وغيرها من المتغيرات الإقتصادية ، وهكذا فإن هذه البيانات المجمعة تؤلف البيانات الخام للعمل الإقتصادي القياسي .

السـبب الرابـع : علـــى الـــرغم مـــن أن الإحصـــاء الرياضـــي (Mathematical Statistics) يوفر عدة أدوات تستخدم في مجالات معينة في الإقتصاد ، فإن المختص بالإقتصاد القياسي غالباً ما يحتاج الى طـــرق خاصـــة للنظر في الطبيعة الأحادية أو الفردانية لمعظم البيانات الإقتصادية ، بمعنـــى إن البيانات لا تخلق بوصفها نتيجة لتجربة مسيطر عليها ، فـــالمختص بالإقتصـــاد القياسي مثل المختص بالأنواء الجوية (Meterologist) عموماً يعتمـــد علـــى بيانات لا يمكن السيطرة عليها مباشرة ، وهكذا فإن بيانات حول الإســـتهلاك أو الدخل أو الإستثمار أو الإدخارات أو الأسعار ... الخ التي تجمع من قبل جهات عامة وخاصة هي ليست بيانات ناتجة عن تجربة .

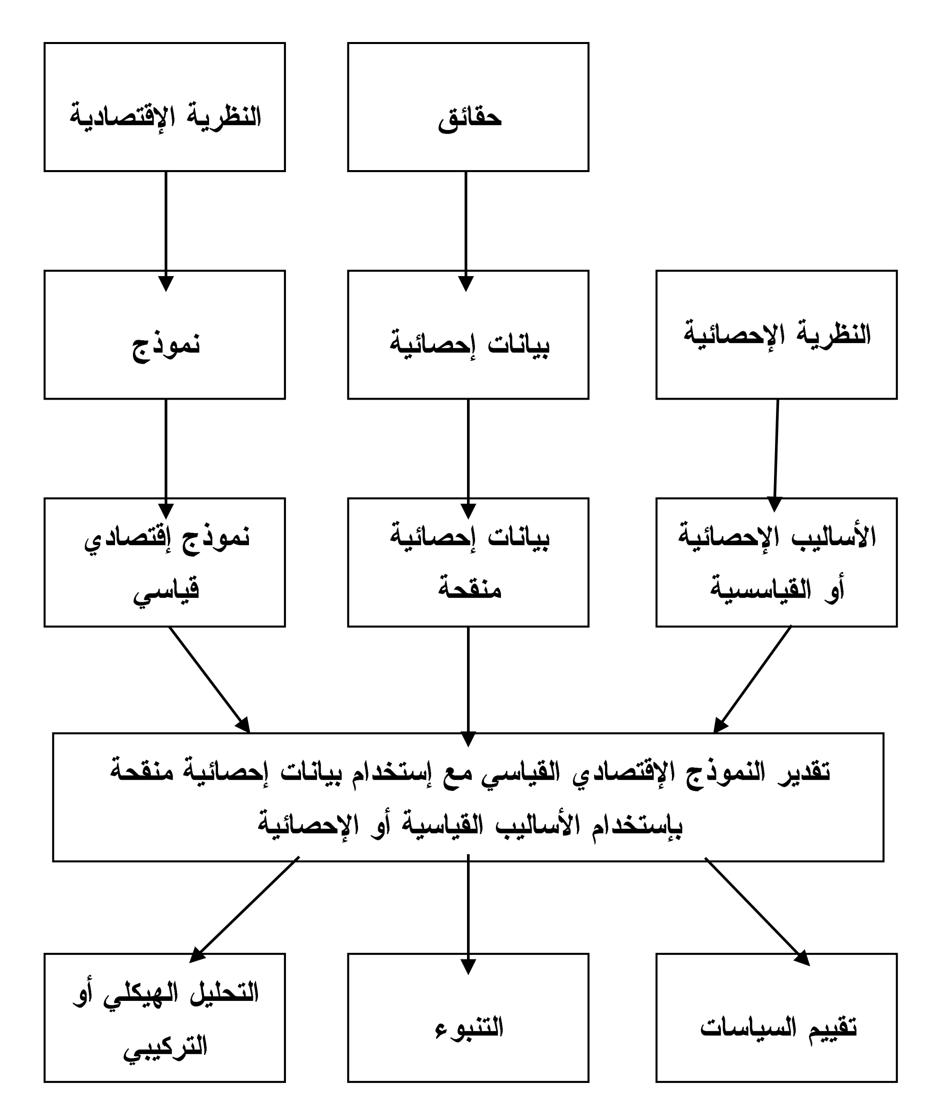
إن المختص بالإقتصاد القياسي يأخذ هذه البيانات كما هـــي معطـــاة ، وهذه تخلق مشكلات خاصة لا يتم التعامل معها ضمن الإحصاء الإقتصـــادي ، والأكثر من هذا فإن مثل هذه البيانات من المحتمل أن تحتوي على أخطاء فـــي

القياس ، مما يدعو المختص بالإقتصاد القياسي لتطوير طرقاً خاصـــة للتحليـــل للتعامل مع مثل تلك الأخطاء في القياس (Errors of Measurment) .

1-3 طبيعة علم الإقتصاد القياسي .

يستخدم علم الإقتصاد القياسي النظرية الإقتصادية والحقائق والأســاليب الإحصائية (Statistical Techniques) .

تعد النظرية الإقتصادية واحدة من الأسس التي يعتمــد عليهــا علــم الإقتصاد القياسي في دراسته للظواهر الإقتصادية ، ويمكن توضيح طبيعة علــم الإقتصاد بالشكل (1-1) الآتي :



شكل (1-1) يوضح طبيعة الإقتصاد

1-4 أغراض الإقتصاد القياسي (Purposes of Econometrics) .

1. التحليل الهيكلي أو التركيبي (Structural Analysis)

عند تقدير النموذج الإقتصادي القياسي يمكننا الحصول على قياسات كمية للعلاقات الإقتصادية المختلفة ، فضلاً عن تسهيل مهمة المقارنة بين النظريات المهمة التي تتعلق بظاهرة إقتصادية واحدة .

يعد التحليل الهيكلي واحداً من أهم الأغراض العلمية لعلـم الإقتصـاد القياسي ، وذلك لتسهيل مهمة فهم الظاهرة الإقتصادية في الحياة الواقعية بطريقة القياس الكمي والإختبار للتأكد من مصداقية العلاقـات بـين معـدل التضـخم (Validating Economic Relationships) والمثال على ذلك قياس العلاقة بين معدل التضخم (Inflation) ومعدل البطالة (Unemployment) ، أو مـا يسمى بـ منحنى فيلبس (Phillips Curve) .

2. التنبوء (Forecasting) .

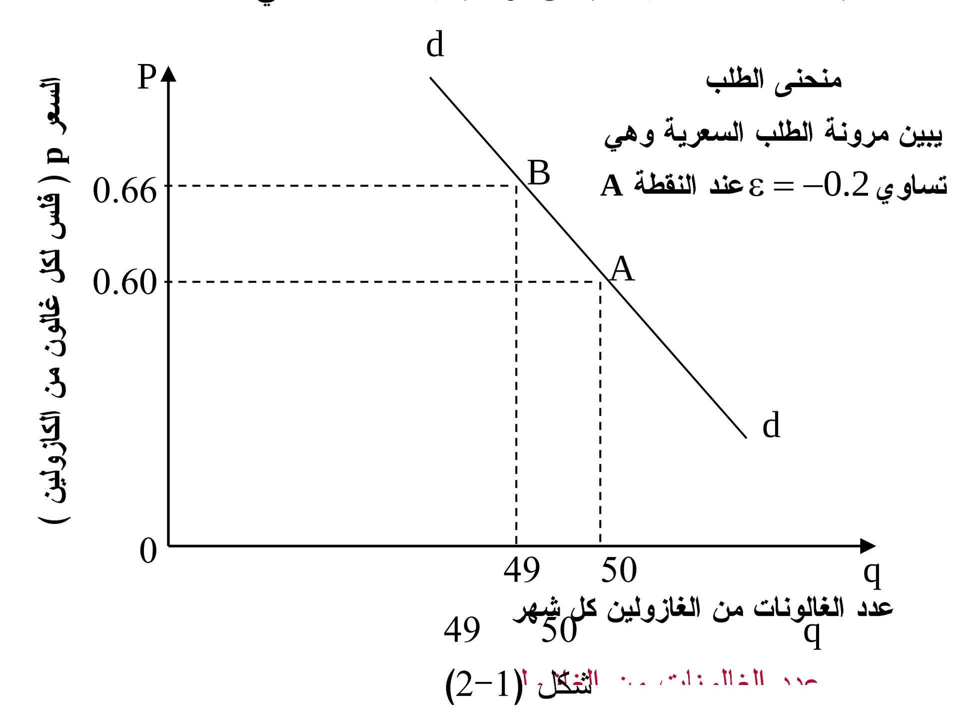
تستخدم نماذج الإقتصاد القياسي للتنبوء عن طريق تقدير تموذج قياسي من أجل التنبوء بالقيم الكمية لمتغيرات معينة لسنوات لاحقة في المستقبل وكمثال على التنبوء شراء المواد الأولية ، وإستخدام عدد من العمال الأضافيين في شركة ما يعتمد على التنبوء بأن المبيعات سـوف تـزداد خـلال السـنتين القادمتين أو أكثر وهكذا .

3. تقييم الساسات الإقتصادية (Policy Evaluation) .

تستخدم نماذج الإقتصاد القياسـي للمفاضـلة والإختيـار بـين السياسات الإقتصادية البديلة ، إن هذه الأغراض للإقتصاد القياسي هي أغراض متداخلة مع بعضها البعض .

مثال (1) منحنى الطلب والمرونة السعرية .

إن واحداً من الأمثلة المهمة في دراسة الإقتصاد القياسي هوتقدير (Estimating) منحنى الطلب لسلعة معينة مثل الكازولين ، إن الطلب على سلعة معينة بالنسبة للمستهلك يمكن توضيحها بالشكل الآتي :



شكل (1−2)

إن الشكل أعلاه يبين العلاقة بين السعر والكمية المستهلكة حيث يمكن وضع هذه العلاقة بصيغة دالة (Function) ، ونقول أن الكمية من الكازولين (q) مقاسة بعدد الغالونات لكل شهر هي دالة للسعر (p) مقاساً بالفلس لكل غالون من الغازولين .

$$q = q(p)$$

إن النقاط كما في الشكل أعلاه للمثال الإفتراضي عن منحنى طلب المستهلك هي (A و B) في نقطة (A) سعر الكازولين كان(0.60) فلساً للغالون

الواحد لذلك ، فإن المستهلك الذي يعبر عنه منحنى الطلب في الشكل أعلاه هــو (dd) سوف يشتري (50) غالوناً من الكازولين لكل شهر .

في نقطة (B) كل الأشياء بقيت على حالها ، ولكن ســعر الكــازولين إرتفع الى (0.66) فلساً للغالون الواحد ، فإن المستهلك ســوف يشــتري (49) غالون في الشهر الواحد ، إن المقياس المفيد الذي يعبر عــن درجــة إســتجابة الكمية المطلوبة من سلعة معينة الى التغيير في سعرها هو مقياس مرونة الطلب السعرية (ε Price elasticity of demand) التي تعرف أنها معدل التغييــر النسبي في الكمية المطلوبة الى التغيير النسبي في السعر .

التغيير النسبي في السعر في أي متغير مثل (z) هو معدل التغير في (z) وليكن (Δz) مقسوماً على المستوى الأساس لــ(z) ، وبعبارة أخرى $\frac{\Delta Z}{Z}$ وهكذا بالطريقة نفسها تكون مرونة الطلب السعرية مكتوبة بالشكل الآتي :

$$\varepsilon = \frac{\Delta q \backslash q}{\Delta p \backslash p} = \frac{p}{q} \cdot \frac{\Delta q}{\Delta p}$$

مرونة الطلب هذه تسمى مرونة الطلب السعرية وهي على نحــو عــام سالبة ، ومن خلال الأرقام الموجودة في الشكل (1-2) ، فإن مرونــة الطلــب السعرية في نقطة (A) تكون تقريباً :

$$\varepsilon = \frac{\Delta q \backslash q}{\Delta p \backslash p} = \frac{\frac{49-50}{50}}{\frac{0.66-0.60}{60}} = \frac{-0.02}{0.10}$$

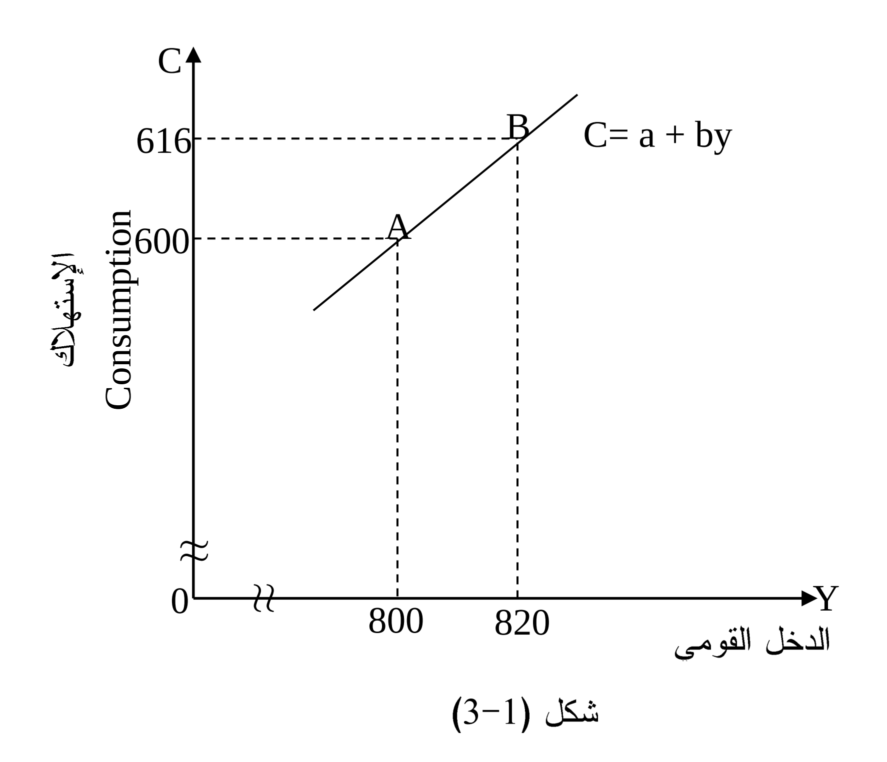
وهكذا في نقطة (A) فإن 10% زيادة في سعر الغالون من الكــازولين سوف تقلل الكمية المطلوبة بنسبة 2% تقريباً . إن تقدير مرونة الطلب السعرية لسلع معينة أو خدمات معينة هو مثال على دراسة الإقتصاد القياسي .

إن هذه الدراسة تجمع بين النظرية الإقتصادية معبر عنها بنموذج منحنى الطلب والحقائق معبر عنها بالأرقام المتعلقة بالسعر والكمية والأسلوب المستخدم هو معادلة المرونة .

إن المقاييس الرقمية ، التي تحصل لدرجة إستجابة الكمية المطلوبة للسعر هي مهمة جداً لأغراض التحليل الهيكلي أو البنيوي ، كما أنها مفيدة لأغراض التنبوء وتقييم السياسات الإقتصادية ، وكمثال التنبوء عن إستيراد السنة القادمة من المواد الخام أو المواد الطبية أو الكيماوية للسنة القادمة .

مثال (2) حول دراسة الإقتصاد القياسي .

دالة الإستهلاك (The Consumption Function) في هذا المثال من دراسة الإقتصاد القياسي تقوم بتقدير دالة الإستهلاك والتي تعد من العناصر الأساسية في بناء النماذج الإقتصادية الكلية التي تقرر الإستهلاك الكلي للإقتصاد الوطني بوصفه دالة في الدخل الكلي .



شكل (1−3)

الشكل (1–3) يوضح دالة إســتهلاك خطيــة وفيهــا قيمــة الإنفــاق الإستهلاكي الكلي (C) بالدينار بوصفه دالة في قيمة الدخل القومي بالدينار (Y) (يعني GNP) ، دالة الإستهلاك على نحو عام هي عبارة عن منحنى صــاعد ، ولكن إنحدارها (Slope) هو أقل من واحد (1) ، بمعنى إن الإضافة الى الدخل تؤدي الى الإضافة للإستهلاك ، ولكن ايضاً تؤدي الى إضافة للإدخارات ، وميل منحنى دالة الإستهلاك هو الميــل الحــدي للإســتهلاك (MPC) (Marginal propensity to consume) ، وهذا الميل من المفروض أن يكون موجبــاً ، ولكن أقل من (1) ، وهكذا في هذه الحالة تكون دالة الإستهلاك الخطية كالآتي :

$$C = a + by \qquad O < b = MPC < 1$$

من النقطتين الموجودتين في الشكل (1–3) فإن (MPC) الميل الحــدي للإستهلاك يمكن تقديره بالطريقة الآتية :

$$\frac{\Delta C}{\Delta Y} = \frac{616-600}{820-800} = 0.8$$

إن الــ (0.8) تتضمن إن (800) فلساً من كل دينار يضاف الى الدخل يذهب الى الإستهلاك .

مرة أخرى نجد أن هذا المثال حول قياس الميــل الحــدي للإســتهلاك (MPC) يمثل دراسة إقتصادية تجمع بين النظرية الإقتصادية (دالة الإستهلاك) والبيانات حول الدخل القومي والإستهلاك الكلي مع اسلوبإقتصادي قياسي عــن طريق معدل التغيير في الإستهلاك على معدل التغيير في الدخل القــومي ، إن هذا القياس مهم في فهم تركيب أو هيكل الإقتصاد الكلي ، وكذلك فــي التنبــوء بمستويات الدخل الكلي أو الإستخدام في المستقبل وكــذلك تحليــل السياســات الإقتصادية المقترحة مثل السياسة النقديــة (Monetary Policy) والسياســة المالية (Fiscal Policy) .

5-1 ما هو النموذج (What is the model) .

النمـــوذج (Model) تعريفـــاً هـــو أي تعبيـــر عـــن ظـــاهرة فعليـــة (Actual phenomenon) ، مثل نظام فعلي أو طريقة معينة ، إذ إن الظاهرة الفعلية أو الحقيقية يعبر عنها بالنموذج من أجل توضيحها والتنبوء بها والسيطرة عليها ، وهذا يذكرنا بأغراض الإقتصـــاد التـــي تضـــمنت التحليـــل الهيكلـــي (Structural Analysis) ، والتنبوء (Forecasting) ، وتقيـــيم السياســـات (Policy Evauation) ، في بعض الأحيان فإن النظام الفعلي يدعى نظام العالم الواقعي (Real – World System) ، وذلك من أجل التمييز بين النظام الفعلي (العالم الحقيقي) ، ونظام النموذج الذي يعبر عنه .

إن فن بناء النماذج هو جزء مكمل لمعظم العلوم بغض النظر عن كونها علوم صرفه (physical) أو علوم إنسانية (Social) وذلك لأن أنظمة العـــالم الحقيقي التي نحن بصددها معقدة على نحو كبير بطبيعتها ، فالنظام قـــد يكـــون الكترون يتحرك في معجل أو أسعار تنظم في أسواق مختلفة ، أو تقرير الدخل القومي ، في هذه الحالات من الأنظمة أو غيرها ، فإن ظاهرة العالم الحقيقيـــة معقدة جداً ، وهذا يتطلب التعامل مع هذه الظاهرة بتعبيرات مبسطة عن طريق النماذج (Models) .

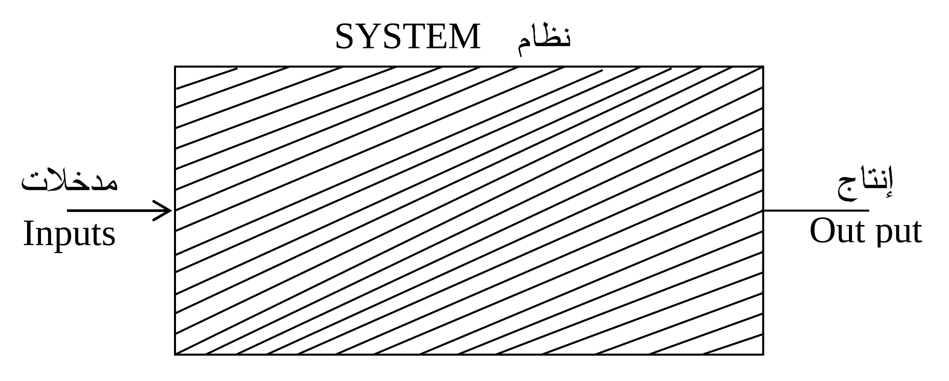
إن أي نموذج يحاول أن يوفق (compromise) بين الحقيقة أو الواقع ، وبين إمكانية التعبير أو إمكانية العمل مع النموذج فضلاً عن إن النموذج يجـــب أن يكون تعبيراً معقولاً عن نظام الواقع أو العالم الحقيقي ، وبهذا المعنى يجـــب أن يكون واقعياً (Realistic) في جمع العوامل الرئيسة للظاهرة المراد التعبير عنها ، ومن جانب آخر يجب أن يكون النموذج ممكـــن العمـــل بـــه أو معـــه (Manageable) في أنه ينتج أو يلقي الأضواء على هذه الظاهرة ، أو ينـــتج معلومات غير ممكن الحصول عليها من الملاحظة أو المراقبة المباشرة للنظـــام الواقعي .

إن الحصول على إمكانية جيدة للعمل مع النموذج يتطلب عادةً طرق مختلفة لتشذيب أو تهذيب النموذج والتي تشتمل على حذف المؤثرات الخارجية كذلك تبسيط العمليات ، إن عملية التشذيب عادةً تجعل النموذج أقل تطابقاً مع الواقع ، ولكن هذه العملية ضرورية للتأكد من أن النظام أو نظام النموذج أصبح من الممكن التعامل معه وإستخدامه على نحو معقول .

وهكذا فإن التوازن الملاءم بين الواقع أو الحقيقة وإمكانية العمل أو إمكانية إستخدام النموذج هو جوهر عملية بناء النموذج الجيد ، ولذلك فإن النموذج الجيد (Good Model) هو النموذج القريب من الواقع وفي الوقت نفسه ممكن العمل به وإستخدامه ، هذا النموذج يحدد العلاقات المتداخلة بين أجزاء النظام للظاهرة بطريقة مفصله ما فيه الكفايه وواضحة ، وذلك للتأكد من إن دراسة النموذج تؤدي الى معلومات تتعلق بنظام العالم الحقيقي ، والنموذج يكون رديئاً عندما يعبر عن الواقع على نحو معقد ، وعندئذ يكون من الصعب التعامل معه وإستخدامه للحصول على المعلومات ، في هذه الحالة فإن مبرر بناء النموذج يصبح معدوماً أو غير وارد بالدرجة الأولى ، والنوع الآخر من النماذج الرديئة يذهب الى تطرف من الجانب الآخر ، أي بمعنى الإبتعاد عن الواقع بمسافة شاسعة ، على نحو يبتعد فيه عن العوامل المهمة في نظام العالم الحقيقي ويحذفها ، وفي هذه الحالة فإن عملية تشذيب النموذج قد ذهبت بعيداً أو تم التركيز عليها على نحو يؤثر على نتائج النموذج وذلك لأن التأثيرات التي يستبعدها النموذج هي مهمة جداً ، إن مثل هذا التطرف على درجة عالية من الخطورة لأنه يوصلنا الى نتائج غير ملائمة ولا تعبر عن حقيقة الحال في الظاهرة الواقعية .

الى هذا المدى نجد أنه من المستحيل أن نصل وعلى نحو دقيق الى كيفية بناء النموذج الجيد ، إذ إن بناء النماذج هو عملية يعد جزءاً منها فناً والجزء الآخر يعد علماً ، إن حالة التطرف تدعى " الصندوق الأسود "

(Black Box) ، في الشكل (1−4) ، إذ لم تبذل أية محاولــة لإعــادة إنتــاج الواقع، أو بعبارة أخرى التعبير عن الواقع على نحو تفصيلي ، في هذه الحالــة فإن النموذج يعبر عن العوامل الداخلة في النظام دون إعتبار النظام نفسه .



الصندوق الأسود (Black – Box) يصف التلفزيون كمثال ويتعــرف على تيار الكهرباء الذي يصل الى التلفزيون ونقاط وعلامات الســيطرة عــن طريق نقطة التشغيل ، وكذلك البرامج التي تقدم على الشاشة ، الصندوق الأسود يعالج الدواخل والخوارج (Inputs and Outputs) دون المحاولة لتحليل كيف ترتبط الدواخل بالخوارج .

إن عملية بناء النموذج عادةً تبدأ مع الصندوق الأسود وبعد ذلك يوضح ما موجود في داخل الصندوق ، إذاً النموذج الأولي هو عبــارة عــن نمــوذج الصندوق الأسود الذي يعالج العوامــل الداخلــه فــي الصــندوق (Inputs) ، والعوامل الخارجة منه (Outputs) فقط وفي بعض الأحيان يســمى النمــوذج الوصفي (Descriptive Model) ولكن متابعة العوامل الداخلة في النظــام أو الصندوق بعد دخولها في النظام ومتابعة النواتج الخارجة قبــل خروجهــا مــن النظام أو الصندوق تؤدي الى نماذج موضحة أكثر وفي النهاية ينــتج نمــوذج تحليلي (Analytic Model) ، أي يصبح الصندوق أبيض (White – Box) الذي يعالج بوضوح التداخل بين العوامل الداخلة والنواتج .

وكمثال فإن نموذج الصندوق الأبيض للتلفزيون يعبـر عنـه بالـدائرة الكهربائية الكاملة ، لذلك فإن عملية بناء النموذج عبارة عن محاولـة مسـتمرة لتشكيل نماذج تحليلية قادرة على تحليل مختلف حالات التداخل للنظام الحقيقي .

6-1 أنواع النماذج (Types of Models) .

ثمة أنواع عدة من النماذج موجودة في العلوم المختلفة ومن بين الأنواع الأكثر أهمية هي النماذج الكلامية المنطقيـة (Verbal \ Logical Models) النمـاذج الفيزيائيـة (Physical Model) ، والنمـاذج الهندسـية (Geometric Models) والنماذج الجبرية (Algebraic Models) .

أولاً : النماذج المنطقية / الكلامية .

تستعمل هذه النماذج التشبيه اللفظي أو الكلامي في العلوم التـي تعتمد على التحقيق (Inquiry) .

إن هذا النوع يستعمل التشبيه مثل الإستعارة أو المجاز والنموذج الناتج بعض الأحيان يدعى الفكرة الشاملة (paradigm) عن حالة معينة ، فـي علـم الإقتصاد هناك مثلاً فكرتين شاملتين الأولى طـورت مـن قبـل آدم سـميث (Adam Smith) وهي فكرة تقسيم العمل التي تتضمن أن كـل فـرد ينجـز الأعمال التي يمتلك فيها مهارة أو ميزة نسبية .

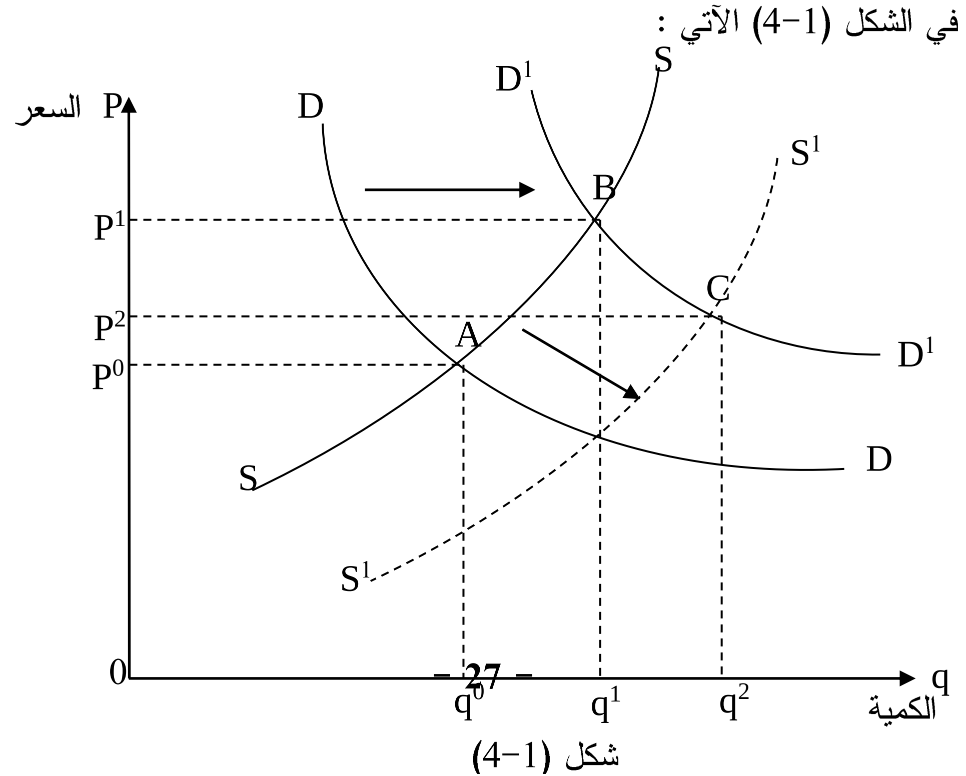
وهكذا يقسم العمل على هذا الأساس مما يؤدي الى زيادة الإنتـاج ، إن هذه الفكرة من الممكن تطبيقها على مستوى الإقتصـاد القـومي أو الـدولي ، إستخدم آدم سميث النموذج الكلامي أو الوصفي ، إذ ناقش هذه الفكرة مشيراً الى معمل إنتاج الدبابيس وتقسيم العمل فيه بوصفه نموذجاً لتوضيح فكـرة تقسـيم العمل .

والفكرة الثانية التـي قـدمها آدم سـميث هـي فكـرة اليـد الخفيـة (Invisible Hand) التي تتضمن إن الأفراد في سعيهم لتحقيق مصالحهم فإنهم

دون أن يشعروا يحققون مصلحة المجتمع في وضع ضمن نظام السوق ، إذ نجد في نظام الأسعار الحره كل وحدة من وحدات الإقتصاد وتقاد في نشاطها عن طريق نظام من علامات التغيير في الأسعار ، وهنا يشبه آدم سميث هذه الحالة بوجود يد خفية توجه كل القرارات التي يصنعها الأفراد بإتجاه الرفاه العام للمجتمع .

ثانياً : النماذج الهندسية (Geometric Models) .

تعد النماذج الهندسية من النماذج ذات الأهمية العظيمة في تطور النظرية الإقتصادية . إذ يتم التعبير عن العلاقات الإقتصادية هندسياً ، ومن أجل أن نبين أهمية هذا النوع من النماذج نشير الى إن معظم كتب النظرية الإقتصادية تستخدم هذا النوع من النماذج والنموذج الهندسي يستخدم المنحنيات والرسوم البيانية الأخرى لتوضيح العلاقات بين المتغيرات ، إن واحداً من الأمثلة المهمة على النموذج الهندسي هو النموذج الذي يستخدم في تقرير السعر في سوق واحدة معزولة ، إذ يمكن الحصول على هذا النموذج عن طريق جمع منحنى الطلب للسوق الواحدة مع منحنى العرض لصناعة واحدة مثل ما موجود في الشكل (1-4) الآتي :



شكل (1-4)

إن منحنى الطلب في السوق هو (DD) يوضح الكمية المطلوبة من السلعة أو الخدمة في مستويات متعددة من الأسعار ، ومنحنى عرض الصناعة (SS) يوضح الكميات المعروضة من السلعة في مستويات بديلة أو متعددة من الأسعار ، تعبر النقطة (A) عن وضع التوازن بين طلب السوق وعرض الصناعة ، بمعنى أنه في نقطة (A) نحصل على التوازن عندما تكون قرارات الشراء من قبل المستهلكين منسجمة مع قرارات البائعين (المنتجين) ، وهكذا فإن الشكل (1−4) يوضح كيفية الوصول الى سعر التوازن وكمية التوازن في سوق واحدة عندما تتقاطع المنحنيات .

أما التغيرات التي تحصل في المتغيرات الأخرى عدا السعر يمكن أخذها في الحساب عن طريق السماح للمنحنيات بالإنتقال أو الحركة ، على سبيل المثال إرتفاع دخل المستهلك أو إنخفاض أسعار السلع المكملة للسلعة في مثالنا ربما يؤدي الى إنتقال منحنى الطلب الى الخارج الى (D^1D^1) ، في حالة مثل هذا الإنتقال في منحنى الطلب يكون المطلوب أكثر من السابق ، عند النقطة (A) سيكون هنالك زيادة في الطلب أكثر من المعروض من السلعة وهذا يؤدي الى وجود طلب غير مشبع (Unsatisfied Demand) ، وهكذا يضع ضغطاً على السعر نحو الإرتفاع ، ونقطة التوازن الجديدة هي (B) .

من جانب العرض ، فإن الحركة أيضاً تشمل حركة منحنى العرض أو إنتقال منحنى العرض ، مثلاً ينتقل منحنى العرض نحو الخارج الى (SS^1) ، وهنا نجد عرضاً أكثر ، إن نقطة التوازن في السعر مع إنتقال المنحنيين نحو الخارج أصبحت نقطة (C) .

ثالثاً : النماذج الجبرية (Algebraic Models)

تعد النماذج الجبرية لأغراض الإقتصاد القياسي مـــن النمـــاذج الأكثر أهمية ، إذ إنها تعبر عن الواقع بوساطة نظام مـــن المعـــادلات ، ومـــن الأمثلة على هذه النماذج ، نموذج دالة الإستهلاك الذي بوساطته يتم تقرير قـــيم التوازن للإستهلاك الكلي والدخل القومي .

$$C = Y (Y) \text{--------------------}(1\text{-}1)$$
$$Y = C + Z \text{---------------------}(1\text{-}2)$$

عندما يكون (Y) هو الدخل القومي (C) الإستهلاك الكلي (Z) الإنفـــاق الحكومي . يتميز النموذج الجبري عن النموذج الهندسي في أن النموذج الجبري يمكننا وبسهولة إضافة متغيرات جديدة الى معادلاته ، ويقرر النموذج الجبـــري قيم متغيرات معينة تدعى المتغيـــرات الداخليـــة (Endogenous Variables) وهي متغيرات معتمدة تقرر قيمتها آنياً بوساطة علاقات النموذج الجبري ، فـــي هذه الحالة فإن الإستهلاك الكلي والدخل القومي هما من المتغيرات الداخليـــة ، والذي نريده هو توضيحها والتنبوء عن قيمها . يحتوي النموذج الجبري أيضـــاً على متغيرات أخرى تدعى المتغيرات الخارجية (Exogenous Variables) ، والتي تقرر قيمها خارج نظام النموذج ولكنها تؤثر على قيم المتغيرات الداخلية . فهي تؤثر على نظام المعادلات ولكنها لا تتأثر بذلك النظام . يحتوي النمـــوذج ايضاً على معلمات (Parameters) معينة ، والتي تقدر على نحو عـــام عـــن طريق أساليب الإقتصاد القياسي بإستخدام الإحصاءات الملاءمة .

رابعاً : النموذج الإقتصادي القياسي (Econometric Model)

يمكن إعتبار نموذج الإقتصاد القياسي نوعاً خاصاً من النمـــاذج الجبريـــة . أولاً إن النمـــوذج الإقتصـــادي القياســـي هـــو نمـــوذج إحتمـــالي (Stochastic Model) يحتوي على واحد أو أكثر من المتغيـــرات العشـــوائية (Random Variables) ، كذلك فهو يعبر عن نظام من العلاقات الإحتماليـــة (Stochastic Relations) بين المتغيرات الداخلة في النظـــام ، وثانيـــاً فـــإن

النموذج الإقتصادي القياسي هو إما نموذج خطي أو نموذج غير خطي ، إن فرضية كون النموذج خطي هي من الفرضيات المهمة جداً ، والمثال على النموذج الإقتصادي القياسي هو النموذج الذي يحتوي على معلمات في المعادلة الآتية :

$$C = a + by \text{ -------------------(1-3)}$$

عندما يكون C = الإستهلاك الكلي ،Y = الدخل الكلي ، b,a = معلمات (b) تعبر عن الميل الحدي للإستهلاك .

من الخصائص المهمة التي يتصف بها النموذج الإقتصادي القياسي هي أن النموذج إحتمالي وهذا النموذج يحتوي على متغيرات عشوائية على عكس النموذج التقريري (Deterministic Model) الذي لا يحتوي على متغيرات عشوائية ، إن المعادلة (1-3) تعطينا مستوى الدخل (Y) ومستوى الإستهلاك (C) مقرر تماماً بقيمة (a + by) ، وهذا غير مقبول أو غير منطقي ، وذلك لإن هناك عوامل أخرى بجانب الدخل يمكن أن تؤثر على الإستهلاك مثل الأسعار ، الذوق ، الثروة .... الخ ، لذلك يجب أن يقدر الإستهلاك (C) عند مستوى معطى من الدخل (Y) كمعدل (a + by) ، وعلى نحو عام فإن الإستهلاك سيكون ضمن قيم عليا وقيم دنيا أي :

$$C = a + by \pm \Delta \text{ --------------(1-4)}$$

وهنا فإن (Є) توضح مستوى أعلى أو أدنى من قيمة المعدل جبرياً أما الطبيعة العشوائية للعلاقة عادةً يعبر عنها في حالة دالة الإستهلاك :

$$C = a + by + \epsilon$$

إن (Є) تعبر عن الخطأ أو الحد العشوائي (Stochastic Disturbance Term) الذي يلعب دور آلية الصدفة (Chance Mechanism) ، إن النموذج الإقتصادي القياسي إما يكون ساكناً أو

حركياً ، فالنموذج الساكن (Static) لا يعتمد على عامل الزمن ، بينما النموذج الحركي فإن الوقت يلعب دوراً أساسياً فيه .

1-7 أنواع العلاقات الإقتصادية

أ- العلاقات السلوكية (Behavioral Relations)

لنأخذ دالة الإستهلاك كمثال وهي :

$$C = a + by \text{ ------------- } (1\text{-}5)$$

إن هذه الدالة تعبر عن علاقة سلوكية ، إذ إنها تصف كيف يتصرف المستهلكون على مستوى المعدل ، آخذين بنظر الإعتبار مجموع مصروفاتهم على السلع والخدمات وبعد معرفة حجوم دخولهم .

ب- العلاقات التعريفية (Definitional Relations)

إن هذه العلاقات تعبر عنها المتطابقات كالآتي :

القيمة بالدينار للسلعة أو الخدمة = الكمية × السعر

التغيير في خزين رأس المال = الإستثمار الصافي

ج- العلاقات الساكنة والحركية (Static & Dynamic Relations)

في هذه العلاقات ، إذا كان كلاً من المتغيرين (C) و(Y) يتعلقان بالسنة نفسها أو الفترة نفسها ، فإن مثل هذه العلاقة تسمى علاقة ساكنة ، ولكن عندما يعتمد الإستهلاك ليس فقط على الدخل في السنة الحالية وإنما يعتمد أيضاً على الدخل في السنة السابقة ، فإننا نستطيع أن نضع المعادلة بالشكل الاتي :

$$C_t = a + b_1 Y_t + b_2 Y_t - 1 \ldots\ldots\ldots\ldots (1\text{-}6)$$

إن المعادلة (1-6) هي مثال يعبر عن علاقة سلوكية حركية وذلك لأنها تحتوي على متغيرات تعود الى تواريخ مختلفة .

1-8 منهجية البحث العلمي في الإقتصاد القياسي

Methodology of Econometrics

لتوضيح منهجية الإقتصاد القياسي لنأخذ النظرية الكنيزية في الإستهلاك يقول كينز (Keynes) :

" إن القانون النفسي الأساسي إن الرجل والمرأة يميلان كقاعدة وعلى مستوى المعدل الى زيادة إستهلاكهم كلما يزداد الدخل ، ولكن ليس بالنسبة نفسها لزيادة الدخل " ، وبإختصار يقول كينز ان الميل الحدي للإستهلاك (MPC) (Marginal Propensity to Consume) وهو معدل التغير في الإستهلاك نتيجة للتغيير بوحدة واحدة مثلاً دينار من الدخل ، وهو أي (MPC) أكبر من الصفر (0) وأقل من (1) $O < MPC < 1$ ، ولإختبار هذه النظرية فإن المختص بالإقتصاد اياسي سيقوم بالإجراءات الآتية :

1-8-1 تحديد النموذج الإقتصادي القياسي

Specification of The Econometric Model

على الرغم من إن كينز يعرض العلاقة الموجبة بين الإستهلاك والدخل ولكنه لم يحدد الشكل الدالي الدقيق للعلاقة بين المتغيرين ، ولغرض التبسيط فإن الإقتصاد الرياضي ربما يقترح علينا الشكل الآتي لدالة الإستهلاك الكينزية :

$$Y = \alpha + \beta X..........(1-7)$$

عندما يكون (Y) = الإنفاق الإستهلاكي ، (X) = الدخل ، $(\beta ، \alpha)$ = ثوابت أو معلمات ، إن معلمة الميل (β) تعبر عن الميل الحدي للإستهلاك (MPC) .

تنص المعادلة (7-1) على إن الإستهلاك يرتبط خطياً بالدخل هي مثال على النموذج الرياضي (Mathematical Model) ، فالنموذج هو عبارة عن منظومة أو مجموعة من المعادلات الرياضية : فإذا إحتوى النموذج على معادلة واحدة ، كما في المثال السابق فإنه يدعى نموذج المعادلة الواحدة (a single equation model) ، بينما إذا إحتوى النموذج على أكثر من معادلة واحدة فإنه يعرف بإنه نموذج متعدد المعادلات

(Multi equation Model) ، أو نـــمـــــــــوذج مـــــــــعادلات آنيـــــة (Simultaneous – equation Model) .

إن النموذج الرياضي الصافي لدالة الإستهـــلاك الذي تمثلـــه المعادلـــة (1−7) محدود الإستخدام من قبل المختص بالإقتصاد القياسي ، وذلك لإن هـــذا النموذج يفترض وجود علاقـــة تامـــة (Exact Relationship) أو علاقـــات تقريرية (Deterministic Relationship) بين الإستهلاك والـــدخل . ولكـــن العلاقات بين المتغيرات الإقتصادية هي عموماً غير تامة (Inexact) . وهكـــذا إذا حصلنا على بيانات حول الإنفاق الإستهلاكي (Y) والدخل تحت التصـــرف (X) (بعد طرح الضريبة) لعينة مؤلفة من (5000) عائلـــة ، ورســـمنا هـــذه البيانات على ورق مربعات مع وضع الإنفـــاق الإســـتهلاكي علـــى الإحـــداثي العمودي والدخل تحت التصرف على الإحداثي الأفقي ، فإننا سوف لن نتوقع أن كل المشاهدات الـــ (5000) أن تقع تماماً على الخط المستقيم للمعادلة (1−7) ، وهذه الحالة تعود الى أنه فضلاً عن الدخل تحت التصرف هناك متغيرات أخرى تؤثر على الإنفاق الإستهلاكي ، على سبيل المثال حجم العائلة ، وأعمار أعضاء العائلة ، وديانة العائلة ...الخ كل هذه المتغيرات تمارس تأثيراً على الإستهلاك العائلي .

وللتعبير عن العلاقات غير التامة (Inexact) بين المتغيرات الإقتصادية فإن المختص بالإقتصاد القياسي سوف يحور دالة الإستهلاك التقريريـــة (1−7) كما يأتي :

$$Y = \alpha + \beta X + U..........(1-8)$$

عندما يكون U = حد الإضطراب أو حد الخطأ (Error Term) وهـــو متغيـــــر عشـــــوائي (Random Variable) أو متغيـــــر إحتمـــــالي (Stochastic Variable) وهو متغير بصفات إحتمالية معروفة جيداً ، إن حد

الإضطراب (U) ربما يعبر عن كل تلك القوى التي تؤثر علـــى الإســـتهلاك ، ولكنها لم تؤخذ في الحساب صراحة (Explicitly) .

إن المعادلة (8-1) تعبر عن مثال للنموذج الإقتصادي القياسي وبصيغة فنية فـــإن المعادلـــة (8-1) هـــي مثـــال علـــى نمـــوذج الإنحـــدار الخطـــي (A Linear Regression Model) ، إن دالة الإستهلاك (8-1) في الإقتصاد القياسي تعبر عن الفرضية التي تقول إن المتغير المعتمد (Y) الإستهلاك يرتبط خطياً بالمتغير التوضـــيحي (Explanatory Variable) (X) الـــدخل ولكـــن العلاقة بين المتغيرين غير تامة ، إنها خاضعة للإنحراف الفردي .

1-8-2 التقدير (Estimation)

بعد تحديد النموذج فإن الخطوة اللاحقة في العمل الإقتصادي القياســـي هي الحصول على تقديرات (Estimates) {قيم رقمية لمعلمات النمـــوذج مـــن البيانات المتوفرة} ، وهذه البيانات يمكن أن تجهز من الإحصاء الإقتصـــادي ، وهكذا في دراسة لدالة الإستهلاك الكينزية المعطاة سابقاً وجد ان $\beta = 0.8$.

إن هذه القيمة الرقمية ليس فقط توفر تقـــديراً رقميـــاً للميـــل الحـــدي للإستهلاك (MPC) ، ولكن أيضاً تدعم فرضية كينـــز التـــي تقـــول إن الـــ (MPC) أقل من (1) ، إن الحصول على تقديرات لـــ β , α ســـيكون مـــن خلال دراسة الإنحدار (تحليل الإنحدار) .

1-8-3 الإثبات (Verification)

الإستدلال الإحصائي (Statistical Inference)

بعد الحصول على تقديرات المعلمـــات β, α فـــإن الخطـــوة اللاحقة هي تطوير معيار ملاءم لإيجاد في ما إذا كانـــت المعلمـــات المقـــدرة بالنموذج تتطابق مع التوقعات من النظرية الإقتصادية . وكما لاحظنا سابقاً توقع كينز أن (MPC) موجبة ولكنها أقل من (1) ، إفترض أنه في دراســـة لدالـــة

الإستهلاك ، وجد أن MPC = 0.9 ، على الرغم من هذا التقدير رقمياً هو أقل من (1) ولكن الباحث لابد أن يعرف في ما إذا كان التقدير دون الواحد ما فيه الكفاية لإقناعنا إن هذه النتيجة ليست نتيجة حصلت بالصدفة في عملية أخذ العينة . بعبارة أخرى ، هل إن هذا التقدير (estimate) إحصائياً أقل من واحد ؟ فإذا كان أقل من (1) فإنه يدعم وجهة نظر كينز ، وإذا كان أكثر من (1) فإنه يدحض رؤية كينز .

إن مثل هذا الإثبات أو الدحض (Refutation) للنظريات الإقتصادية على أساس الدلائل التجريبية يتأسس على فرع من فروع النظرية الإحصائية يدعى الإستدلال الإحصائي (Statistical Inference) أو إختبار الفرضيات (Hypothesis Testing) .

4-8-1 التنبوء (Forecasting)

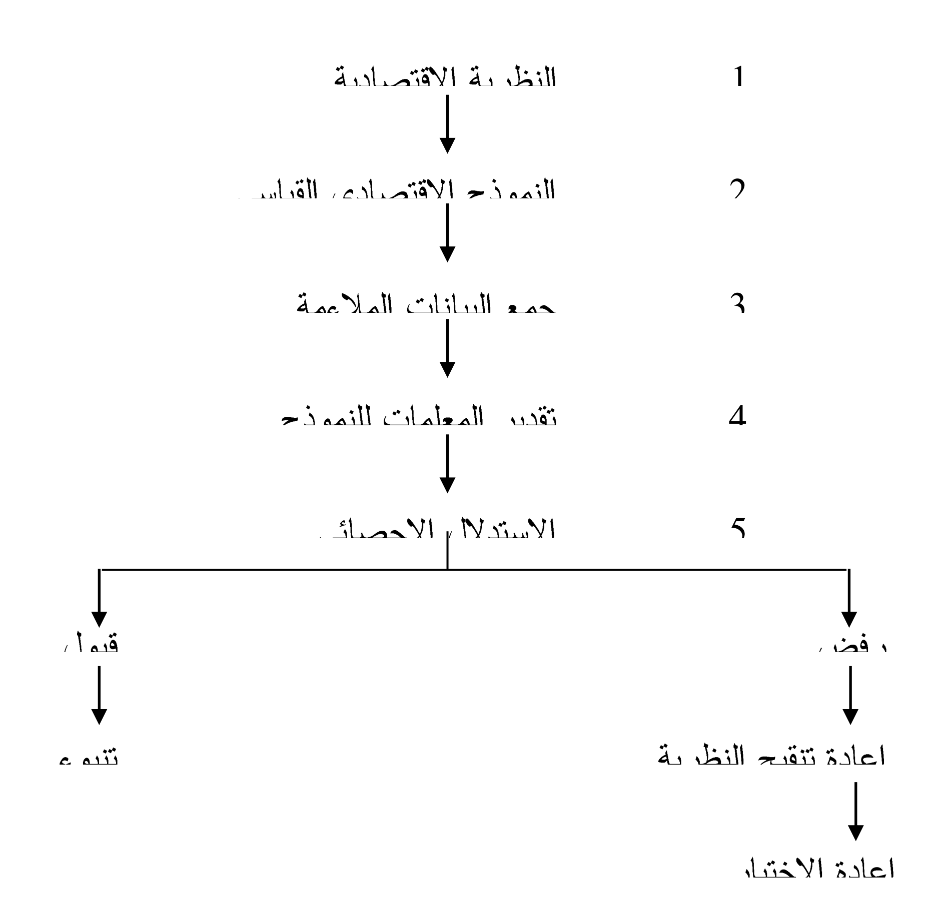
غالباً ما يستخدم النموذج الإقتصادي القياسي المقدر لأغراض التنبوء بقيم مستقبلية للمتغير المعتمد على أساس قيم معروفة أو متوقعة في المستقبل للمتغير المستقل ، مثلاً إفترض إن الحكومة تقوم بتخفيض في ضريبة الدخل الفردية أو الشخصية من أجل حفز الإقتصاد القومي ، ما هو التأثير الذي سوف تحدثه هذه السياسة على الإنفاق الإستهلاكي ، ولذلك على الإستخدام والدخل ، وكما تبين نظرية الإقتصاد الكلي فإن التغيير في الإنفاق الإستهلاكي الذي ينتج بعد تغيير بمقدار دينار واحد في الدخل يكون معطى بوساطة مضاعف الإستهلاك (Consumption Multiplier) (M) الذي يعرف كما يأتي :

$$M = \frac{1}{1-MPC}$$

إذا كان MPC = 0.8 فإن (M) سيكون (5) ، بمعنى إذا إزداد الدخل بدينار واحد فإن ذلك سيقود الى زيادة خمسة مرات (Fivefold) في الإنفاق الإستهلاكي ، إن القيمة الحاسمة في هذا الحساب هو مضاعف الإستهلاك الذي

يعتمد على قيمة (MPC) . وهكذا فإن التقـدير الكمـي لـــ (MPC) يــوفر معلومات قيمة لأغراض السياسة ، فــإذا عرفنــا (MPC) نســتطيع التنبــوء بالإستهلاك في المستقبل بعد تغييرات في سياسة الحكومة المالية .

وهكذا فإن التحقيق الإقتصادي القياسي عموماً يكون كما يأتي :



1-9 أنواع الإقتصاد القياسي

عموماً فإن الإقتصاد القياسي يمكن أو ربما يقسيم الى فئتين :

1- الإقتصاد القياسي النظري (Theoretical Econometrics)

2- الإقتصاد القياسي التطبيقي (Applied Econometrics)

الأول يهتم بتطوير الطرق الملاءمة لقياس العلاقات الإقتصادية المحددة بوساطة نماذج الإقتصاد القياسي ، وفي هذا الإعتبار فإن الإقتصاد القياسي يعتمد على نحو كبير على الإحصاء الرياضي (Mathematical Statistics) .

أما في الإقتصاد القياسي التطبيقي فإننا نستعمل أدوات الإقتصاد القياسي النظري لدراسة حقول خاصة في الإقتصاد مثل دالة الإنتاج ودالـــة الإســـتهلاك ودالة الإستثمار ودوال العرض والطلب .... الخ .

الفصل الثاني

نظرية الإرتباط

نظرية الإرتباط

2-1 مقدمة

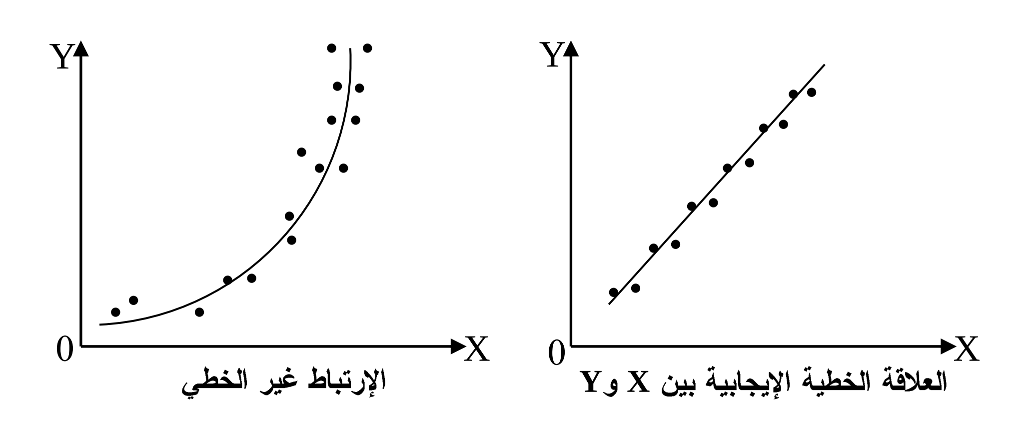
ثمة طرق مختلفة لقياس العلاقات الموجودة بين المتغيرات الإقتصادية ومن أبسط هذه الطرق هي طريقة تحليل الإرتباط (Correlation Analysis) وطريقة تحليل الإنحدار (Regression Analysis) ، يجب أن نبدأ من تحليل الإرتباط وذلك على الرغم من أن تحليل الإرتباط يعاني من محـــددات خطيــرة عديدة ، فضلاً عن أنه يلقي ضوءاً بسيطاً على طبيعة العلاقة الموجــودة بـــين المتغيرات ، ولكنه من ناحية أخرى يجعل القاريء أو الطالب على معرفة جيدة بمعامل الإرتباط الذي يعد من المقاييس الإحصائية المهمة في تحليل الإنحدار .

يعرف الإرتباط بأنه درجة العلاقة الموجودة بين متغيران أو أكثر ، إن درجـــة العلاقـــة الموجـــودة بـــين متغيـــرين تســـمى إرتبـــاط بســـيط (Simple Correlation) , ودرجة العلاقة التي تربط ثلاث متغيرات أو أكثر تسمى إرتباط متعدد (Multiple Correlation) ، سيتم التركيز في هذا الفصل على الإرتباط البسيط فقط ، تاركين مناقشة الإرتباط المتعدد الى فصل آخر بعد مناقشة تحليل الإنحدار (والحقيقة أن معامل الإرتباط المتعدد لا يمكن تفسيره أو دراسته من دون دراسة تحليل الإنحدار المتعدد) .

ربما يكون الإرتباط خطياً عندما تكون نقاط (Y،X) في الشكل البيـــاني أو شكل الإنتشار (Scatter Diagram) تبدو متركزة قرب خط مســـتقيم ، أو غير خطي (Non-Linear) عندما تكون نقاط (Y،X) تبدو واقعة قرب منحنى إن متغيرين إثنين ربما يكون لهما إرتباط إيجابي أو إرتباط سلبي أو ربما ليس لهما إرتباط (Uncorrlated) ، وهذا الشيء نفسه يصح بالنسبة للعلاقة الخطية وغير الخطية .

2-1-1 الإرتباط الإيجابي (Positive Correlation) .

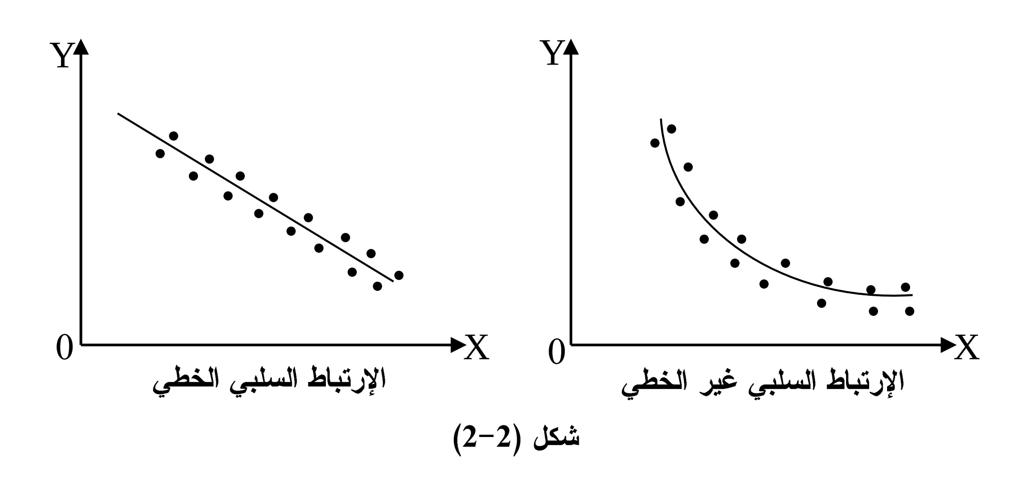
إذا كان هناك متغيرين مرتبطين إيجابياً ، فإن ذلك يعني أن كـلاً مـن المتغيرين يميلان الى التغيير في الإتجاه نفسه أي أن المتغيرين يميلان للتغيير معاً وفي الإتجاه نفسه ، وهذا يعني أنهما يميلان الى الزيادة والنقصان معاً ، إن مثل هذا الإرتباط الإيجابي تفترضه النظرية الإقتصادية في علاقة عرض السلعة من الكميات مع سعر السلعة في السوق ، عندما يـزداد السـعر فـإن الكميـة المعروضة تزداد وبالعكس عندما ينخفض السعر فإن الكمية المعروضة تنخفض أيضاً ، وكما موضح في الشكل الآتي (2-1) :



العلاقة الخطية الإيجابية بين X و Y الإرتباط غير الخطي

2-1-2 الإرتباط السلبي (Negative Correlation) .

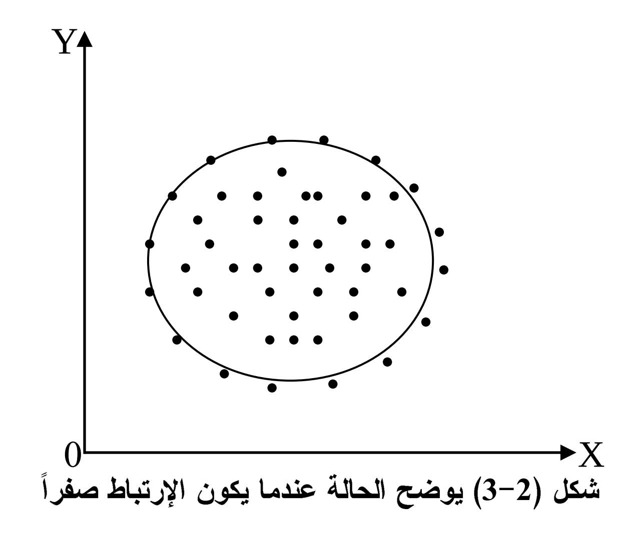
إن متغيرين يقال أنهما مرتبطان سلبياً إذا كانا يميلان للتغيير بإتجـاه معاكس : أي عندما يزداد (X) فإن (Y) ينخفض والعكس بالعكس ، مثلاً الكمية المطلوبة من السلعة وسعرها يرتبطان سلبياً ، فعندما يزداد السعر فإن الطلـب على السلعة ينخفض ، وعندما ينخفض السعر فإن الطلب على تلك السلعة يزداد والشكل البياني الآتي (2-2) يوضح أن نقاط (X و Y) تتمركـز حـول خـط مستقيم أو منحنى بميل سالب ، فإذا كانت كل نقاط (X و Y) تقع على الخـط المستقيم أو المنحنى فإن الإرتباط هو إرتباط سلبي تام .



شكل (2-2)

2-1-3 عدم الإرتباط (No Correlation, or Zero Correlation) .

إن متغيرين غير مرتبطين مع بعضهما عندما يميلان للتغيير مع عدم الإرتباط ببعضهما ، وكما توضح في الشكل البياني (2-3) الآتي أو شكل الإنتشار (Scatter Diagram) .



شكل (2-3) يوضح الحالة عندما يكون الإرتباط صفراً

إن النقاط لـ(X,Y) منتشرة على سطح المستوى لـ (X,Y) مثلاً نقول إن الإرتباط صفراً بين وزن الطلبة ولون شعرهم .

2-2 قياس الإرتباط الخطي (Measure of Linear Correlation) .

2-2-1 معامل الإرتباط الحقيقي (إرتباط المجتمع الإحصائي The population correlation Coefficient (P) ومعامل الإرتباط المقدر للعينة Sample Estimate (r) .

على ضوء المناقشة السابقة يبدو أنه بإمكاننا أن نقرر نوع الإرتباط بين متغيرين بوساطة المشاهدة المباشرة للرسم أو الشكل البياني ، فضلاً عن أن الرسم البياني يوضح قوة العلاقة بين متغيرين ، فإذا وقعت النقاط قريبة من الخط المستقيم ، فإن الإرتباط يكون قوياً ، من ناحية أخرى فإن التشتت العظيم (dispersion) في النقاط حول الخط يتضمن إرتباطاً ضعيفاً ، من الملاحظ أن فحص الشكل البياني يعطينا فكرة مبسطة أو صورة تقريبية للعلاقة بين المتغيرين (x و y) .

من أجل الحصول على قياس دقيق لدرجة العلاقة أو الإرتباط بين المتغير (X) والمتغير (Y) نستعمل معامل يدعى معامل الإرتباط (The Correlation Coefficient) ، وعادةً يرمز له بالحرف الأغريقي (P) إن الحرف (P) يشير الى كل القيم أو المجتمع الإحصائي (population) للمتغيرين (X) و(Y) ، إن تقدير (P) من أية عينة إحصائية (Sample) يدعى بـ (r) ، مثلاً إذا كنا نقيس الإرتباط بين (X) و(Y) فإن معامل إرتباط المشاهدة الكلية أو المجتمع الإحصائي يعبر عنه بـ (P_{XY}) وتقديره من العينة الإحصائية يعبر عنه بـ (r_{XY}) معامل الإرتباط في العينة ، يمكن كتابته على النحو الآتي :

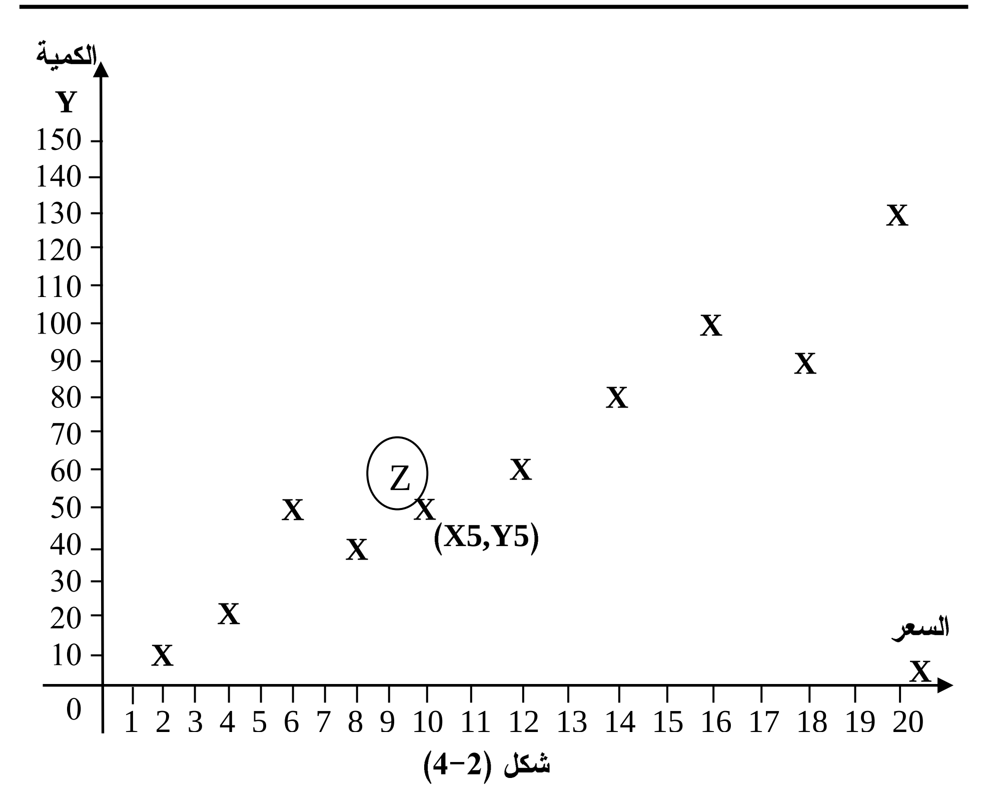
$$rXY = \frac{\sum(xi-\bar{x})(yi-\bar{y})}{\sqrt{\sum(xi-\bar{x})2}\sqrt{\sum(yi-\bar{y})2}}$$

سوف نستخدم مثالاً بسيطاً من نظريـــة العـــرض (Supply Theory) توضح النظرية الإقتصادية أن الكمية المعروضة من السلعة في السوق تعتمـــد على سعرها مع بقاء الأشياء الأخرى ثابتة (Cetris paribus) عنـــدما يـــزداد السعر فإن الكمية المعروضة تزداد وبالعكس ، وعندما يـــنخفض الســـعر فـــإن المنتجين يعرضون كميات أصغر من سلعهم للبيع ، بعبارة أخرى إن النظريـــة الإقتصادية ترى إن السعر (X) والكمية المعروضة (Y) مرتبطان إيجابياً .

إن مشكلتنا هنا أن نعرف المقياس الذي سنستخدمه لتقدير الإرتباط بـــين السعر (X) والكمية المعروضة (Y) ، واجبنا الأول هـــو جمـــع المشـــاهدات (الأسعار والكميات المعروضة خلال مدة معينة من الزمن) ، بعـــض البيانـــات الإفتراضية تبدو في الجدول (2-1) الآتي :

السعر بالدينار Xi	الكميات المعروضة بالكغم Yi	المدة الزمنية بالأيام
2	10	1
4	20	2
6	50	3
8	40	4
10	50	5
12	60	6
14	80	7
16	90	8
18	90	9
20	120	10
$\sum_{i}^{n} Xi = 110$	$\sum_{i}^{n} Yi = 610$	N=10

عندما نضع هذه البيانات بصيغة شكل الإنتشار (Scatter Diagram) نحصل على الشكل البياني (2-4) ، وذلك من أجل أن نعرف هل أن العلاقـــة خطية (Linear) أو غير خطية (Non Linear) .



شكل (2−4)

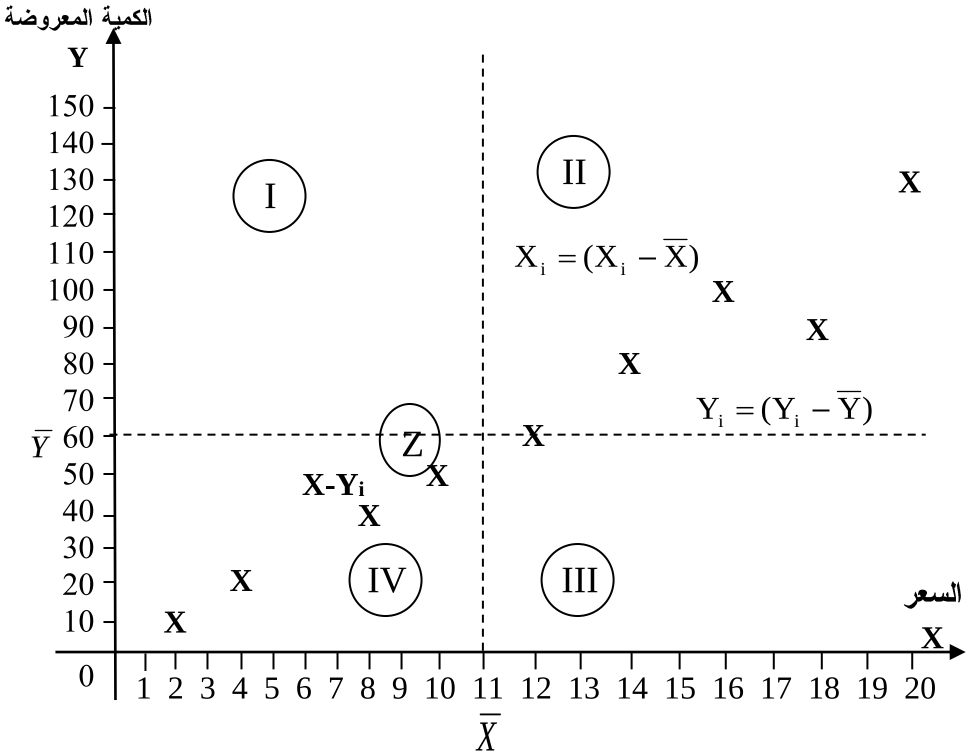
إن كل نقطة في شكل الإنتشار تعبر عن زوج من السعر والكمية في مدة زمنية معينة على سبيل المثال النقطة (Z) تعبر عن زوج (X_5,Y_5) وهي النقطة التي يكون فيها السعر عشرة دنانير والكمية المعروضة (50)كغم خلال المدة الخامسة ، ومن النظر الى شكل الإنتشار ، نرى أن النقاط تميل الى الإستقطاب حول خط ذو ميل موجب ، وهذا يثبت أن هناك إرتباط خطي إيجابي بين السعر والكمية ، من أجل إيجاد المقياس التام للإرتباط نعمل كما يأتي :

1− نحسب كمية الوسط الحسابي (Mean) للمتغيرات .

$$\overline{Y} = \frac{\sum Yi}{n} = \frac{610}{10} = 61$$

$$\overline{X} = \frac{\sum Xi}{n} = \frac{110}{10} = 11$$

2- نرسم الأعمدة $\overline{XX}$ ، $\overline{YY}$ من الوسط الحسابي $\bar{X}$ والوسط الحسابي $\bar{Y}$ وهكذا نقسم المساحة الى أربعة أرباع IV,III,II,I في الشكل (2-5) الآتي :



3- بعدها نأخذ إنحرافات قيم X و Y عن الوسط الحسابي لـــــ X و Y والإشارة الى هذا الإختلاف بالحروف الصغيرة وكما يأتي :

$$xi = (Xi - \bar{X})$$

$$yi = (Yi - \bar{Y})$$

إن نتيجة هذه العملية في إختيار الإنحرافات لقيم المتغيرات X و Y عن الوسط الحسابي ، يمكن أن تزودنا بمقياس للإرتباط بين المتغير (X) والمتغيــر (Y) .

(a) في الربع الثاني (II)، والربع الرابع (IV) فإن :

الناتج $(Xi-\bar{X})(Yi-\bar{Y})=xiyi$ موجباً ، وذلك لإن للإنحرافين $(Xi-\bar{X})$ و $(Yi-\bar{y})$ الإشارة نفسها إما موجبة أو سالبة .

(b) في الربع الأول (I)، والربع الثالث (III) فإن ناتج الإنحرافات $(Xi-\bar{X})(Yi-\bar{Y})=xiyi$ سالباً لإن للإنحرافات إشارات الواحدة عكس الأخرى.

وهكذا إذا كان أغلب المشاهدات (البيانات) تقع في الربع الثاني (II) والربع الرابع (IV) ، فإن الإرتباط بين (X,Y) يكون موجباً ،وإذا كانت معظم النقاط للسعر والكمية تقع في الربع الأول (I) والربع الثالث (III) ، فإن الإرتباط بين (Y,X) سيكون سالباً ، وإذا كانت المشاهدات أو النقاط منتشرة على نحو عشوائي في كل الأرباع أو الأجزاء الأربعة ، فإن القيم الموجبة والقيم السالبة لـ(xiyi) ستلغي بعضها البعض ومجموع النتائج يمثل الإقتراب من الصفر ، فإذا كان مجموع ناتج الإنحرافات للمتغيران (X)(Y) عن وسطهما الحسابي هو قيمة موجبة فإن الإرتباط بين (X) و(Y) سيكون موجباً ، بينما إذا كان مجموع ناتج الإنحرافات سالباً فإن الإرتباط بين (X),(Y) سيكون سالباً :

$$\sum_{i}^{n}(Xi-\bar{X})(Yi-\bar{Y})=\sum_{i}^{n}xiyi\succ 0$$

الإرتباط بين (X),(Y) موجباً .

$$\sum_{i}^{n}(Xi-\bar{X})(Yi-\bar{Y})=\sum_{i}^{n}xiyi\prec 0$$

الإرتباط بين (X),(Y) سالباً .

وهكذا فإن ناتج مجموع الإنحرافات $(\sum xiyi)$ يزودنا بمقياس للعلاقة (X),(Y) ولكن هذا القياس له خللين :

<u>الأول</u> : إن هذا المقياس يؤثر عليه عدد المشاهدات الداخلة . كلما كان عدد المشاهدات كبيراً كلما كان عدد الناتج أكبر ولذلك فإن مجموع $(\sum xiyi)$

سيكون مختلفاً . وهكذا إذا كان (X),(Y) مرتبطان إيجابياً ، فإن أي زيادة في عدد المشاهدات سوف تجعل الإرتباط يبدو قوياً بدون أن يكون ذلك صحيحاً بالضرورة .

الثاني : إن المجموع ($\sum xiyi$) يتأثر بوحدات القياس للمتغيرات (X),(Y) مثلاً الإرتباط في المثال السابق سوف يبدو أكبر إذا كان العرض مقاساً بالكيلوغرام والسعر بالفلس حتى لو كانت المشاهدات أي العدد نفسه من المشاهدات .

لتصحيح الخلل الأول نقوم بتقسيم المجموع ($\sum xiyi$) على عدد المشاهدات(n) .

$$\frac{\sum xiyi}{n} = Sxiyi$$

إن Sxyهو التباين المشترك (Covariance) بين (X)و(Y) من الواضح انه مقياس أفضل للإرتباط من المجموع البسيط ($\sum xiyi$) وذلك لأنه سوف يتغير مباشرة مع عدد المشاهدات في العينة (Sample) ولكن مع هذا يبقى المقياس يعاني من خلل لأنه يتأثر بالوحدات التي تقاس بها المتغيرات (X),(Y) ومن أجل تصحيح هذا الخلل نقوم بتقسيم التباين المشترك (Covariance) على الإنحرافات المعيارية (Standard deviations) للمتغيرات التي تقاس بالوحدة نفسها التي تقاس بها المتغيرات التي تكون أو تصبح النسبة رقم صافي مستقل عن أي تغيير في وحدات قياس (X),(Y) ، إن النسبة الناتجة هي معامل إرتباط العينة (r) الذي هو تقدير لمعامل إرتباط المشاهدات أو المجتمع الإحصائي (P) .

$$r = \frac{\sum_{i}^{n} xiyi}{nSxSy} = \frac{S_{YX}}{S_X S_Y}$$

S_{xy} = (Y),(X)التباين المشترك لـــ $= \frac{\sum XIYI}{n}$

S_x = (X)الإنحراف المعياري لـــ

$$S_X = \sqrt{\frac{\sum (XI - \overline{X})^2}{n}} = \sqrt{\frac{\sum xi2}{n}}$$

S_y= (Y)الإنحراف المعياري لـــ

$$SY = \sqrt{\frac{\sum (Yi - \overline{Y})^2}{n}} = \sqrt{\frac{\sum yi^2}{n}}$$

نعوض عن قيم S_Y, S_X, S_{XY} في معادلة (r) نجد :

$$r = \frac{\sum xiyi}{n\sqrt{(\frac{\sum xi^2}{n})(\frac{\sum yi^2}{n})}} = \frac{\sum xiyi}{\sqrt{(\sum xi^2)(\sum yi^2)}} \quad ....(1-2)$$

هذه المعادلة وضعت بصـــيغة الإنحرافــات للمتغيــرات(X),(Y)عـــن وسطيهما الحسابي ، أما إذا أردنا أن نستخدم القيم الفعليـــة للمشـــاهدات فإننـــا نستخدم الشكل الآتي للمعادلة :

$$r = \frac{n\sum(XiYi) - (\sum Xi)(\sum Yi)}{\sqrt{n\sum Xi^2 - (\sum Xi)^2}\sqrt{n\sum Yi^2 - (\sum Yi)^2}} \quad ....(2-2)$$

هذا الشكل للمعادلة مشتق من المعادلة السابقة عن طريـــق التحـــويلات الآتية :

$$r = \frac{\sum xiyi}{\sqrt{\sum xi^2}\sqrt{\sum yi^2}}$$

1- البسط (Numerator) يمكن توسيعه كما يأتي :

$$\sum xiyi = \sum (X - \overline{X})(Y - \overline{Y}) = \sum (XY - Y\overline{X} - X\overline{Y} + \overline{X}\overline{Y})$$

$$= \sum XY - \overline{X}\sum Y - \overline{Y}\sum X + n\overline{X}\overline{Y}$$

$$= \sum XY - \frac{\sum X}{n}\sum Y - \frac{\sum Y}{n}\sum X + n\frac{\sum X}{n}\frac{\sum Y}{n}$$

$$= \frac{n\sum XY - \sum X\sum Y}{n} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots(3-2)$$

2- من المقام في المعادلة (2-1) حصلنا على :

$$\sum xi^2 = \sum (Xi - \overline{X})^2 = \sum (Xi^2 - 2X\overline{X} + \overline{X}^2)$$

$$= \sum Xi^2 - 2\overline{X}\sum X + n\overline{X}^2$$

$$= \sum Xi^2 - 2\overline{X}\sum X + nX\frac{\sum Xi}{n}$$

$$= \sum Xi^2 - 2\overline{X}\sum Xi + \overline{X}\sum X$$

$$= \sum Xi^2 - \overline{X}\sum Xi$$

$$= \sum X^2 - \frac{\sum Xi}{n}\sum Xi$$

$$= \sum X^2 - \frac{(\sum X)^2}{n} \ldots\ldots\ldots(4-2)$$

$$= \frac{n\sum Xi^2 - (\sum Xi)^2}{n} \ldots\ldots\ (5-2)$$

وبالطريقة نفسها نستطيع أن نجد قيمة $\sum yi^2$ للمتغير Y

$$\sum yi^2 = \frac{n\sum Y^2 - (\sum Y)^2}{n} \quad \ldots\ldots\ldots\ldots(6-2)$$

3- نعوض عن المعادلة (2-3)(2-4)(2-6) في المعادلة (2-1) نجد أن :

$$r = \frac{\left\{n\sum XY - (\sum X)(\sum Y)\right\} \ /n}{\sqrt{\frac{\left\{n\sum X^2(\sum X)^2\right\}}{n}}\sqrt{\frac{\left\{n\sum Y^2-(\sum Y)^2\right\}}{n}}}$$

$$r = \frac{n\sum XY - (\sum X)(\sum Y)}{\sqrt{n\sum X^2-(\sum X)^2}\sqrt{n\sum Y^2-(\sum Y)^2}}$$

إن القانون المذكور أعلاه لمعامل الإرتباط له فائدتين :

1- القانون متشابه بالنسبة الى (X)و(Y) وهذا يعني أن $\mathbf{r_{xy}=r_{yx}}$

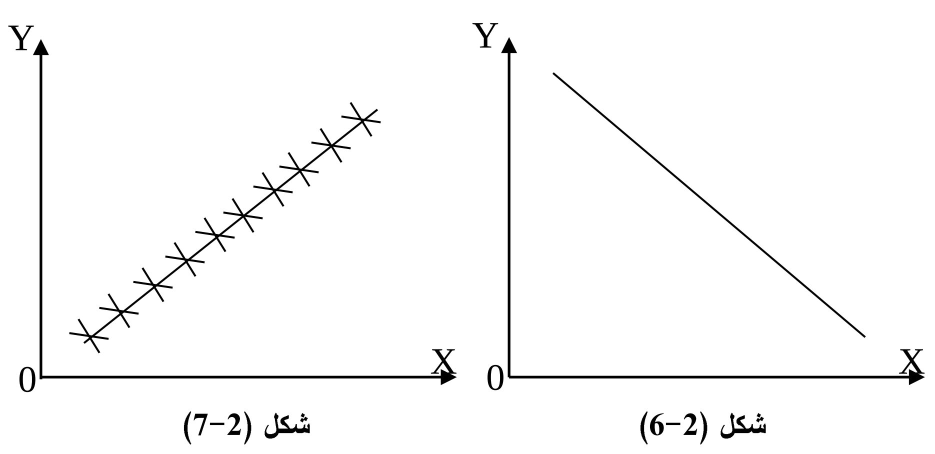
2- القانون أعلاه يطبق فقط في حالات العلاقات الخطية .

2-2-2 القيم الرقمية لمعامل الإرتباط

إن معامل الإرتباط هو مقياس لدرجة التغيير المتزامن للمتغيرات مع بعضها البعض (Covariability) (X) و(Y) .

إن القيم التي يتخذها معامل الإرتباط بين -1 الى +1 عندما يكون (r) موجباً فإن المتغير (X) و(Y) يتغيرات بالإتجاه نفسه زيادة أو نقصاناً معاً ، وعندما يكون (r= +1) فهذا يعني أو يتضمن أن هناك إرتباط تام إيجابي بين (X) و(Y) .

وبالرسم البياني فإن كل المشاهدات عن (X) و(Y) تقع على خط مستقيم مع ميل موجب كما في الشكل (2-7) عندما يكون معامل الإرتباط (r) سالباً فإن (X) و(Y) يتحركان بإتجاهات متعاكسة ، فإذا كان r = -1 فإن هذا يعني أن هناك علاقة إرتباط سالبة تامة موجودة بين المتغيرين (X) و(Y) وبالشكل البياني كل المشاهدات عن(X) و(Y) تقع على خط مستقيم مع ميل سالب كما في الشكل (2-6) وعندما يكون (r) مساوياً للصفر فإن المتغيرين غير مرتبطين (Uncorrelated) .



شكل (2−6) شكل (2−7)

في التطبيق العملي سوف لا نلاحظ حالة الإرتباط التام أو حالة عدم الإرتباط ، عادة فإن (r) يأخذ قيماً بين الصفر والواحد ، كلما كانت قيمة (r) قريبة من قيمة (1) فإن درجة الإرتباط تكون كبيرة وإن النقاط قريبة أو تقترب من الخط المستقيم في الشكل البياني ، من ناحية أخرى كلما كانت النقاط مبعثرة في الشكل البياني كلما إقترب الإرتباط بين المتغيرين من الصفر .

نقول أن (r) هو معامل الإرتباط المقدر من العينة لمعامل الإرتباط للمجتمع الإحصائي (P)(population Correlation) . إن (r) بوصفه تقديراً إحصائياً حتماً معرض لبعض الأخطاء ويجب أن يختبر من ناحية درجة الإعتماد إحصائياً .

مثال : لنفترض أننا نريد أن نحسب معامل الإرتباط بين المتغير (Y)(الكمية المعروضة) و(X)(السعر) مع المشاهدات العطاة في الجدول (1−2):

حساب معامل الإرتباط (r) بإستعمال الإنحرافات عن الوسط الحسابي

$$r = \frac{\sum xiyi}{\sqrt{\sum xi^2}\sqrt{\sum yi^2}}$$

في هذا المثال نريد حساب قيم $\sum xiyi$ ، $\sum xi^2$ ، $\sum yi^2$ الموجودة في القـــــانون أعـــــلاه : $\sum xi^2 = 330$ و $\sum xiyi = 1810$ و $\sum yi^2 = 10490$
وبالتعويض نجد أن :

$$r = \frac{1810}{\sqrt{330}\ \sqrt{10490}} = 0.975$$

حساب معامل الإرتباط بإستخدام أو إستعمال البيانات الفعلية :

$$r = \frac{n\sum XiYi - \sum Xi \sum Yi}{\sqrt{n\sum X^2 - (\sum X)^2}\ \sqrt{n\sum Y^2 - (\sum Y)^2}}$$

من القانون نجد أننا بحاجة الى حساب القيم الآتية :

$$\sum Xi^2 \quad \sum Yi^2 \quad \sum Xi \quad \sum Yi \quad (\sum Xi)^2 \quad (\sum Yi)^2 \quad \sum XiYi$$

$$r = \frac{(10)(8.520) - (610)(110)}{\sqrt{(10)(1540) - 12100}\ \sqrt{(10)(47.700) - 372100}}$$

$$r = 0.975$$

3-2 معامل إرتباط الرتب (The Rank Correlation Coefficient).

إن معامل الإرتباط الخطي الذي قمنا بدراسته في النقطة السابقة يعتمـــد على فرضية أن المتغيرات الداخلة في العلاقة هي متغيـــرات كميـــة ، ولـــدينا بيانات دقيقة من أجل قياس تلك المتغيرات . ولكن في حالات عدة ربما تكـــون المتغيرات نوعية (Qualitative Variables) ولذلك لا يمكن قياسها بالأرقام (numerically) ، مثلاً المهنة والتعليم والتفضيلات لنوع معين وغيرها مـــن المتغيرات النوعية ، فضلاً عن أنه في حالات عدة فإن قيماً دقيقة للمتغيـــرات ربما لا تكون موجودة ، لذلك فإنه من المستحيل حساب قيمة معامل الإرتبـــاط

مع القانون الذي ذكرناه في النقطة السابقة ، في مثل هذه الحالة من الممكن أن نستخدم مقياس إحصائي آخر وهو معامل إرتباط الرتب أو ما يسمى بـــ معامل إرتباط سيبرمان (Spearman Correlation Coefficient) نقــوم بترتيــب المشاهدات او البيانات في تتابع محدد مثلاً حسب الحجم أو الأهمية .... الــخ مستعملين الأرقام 1،2،3.....n بعبارة أخرى نحن نعطــي رتــب (Ranks) للبيانات الإحصائية ونقيس العلاقة بين الرتب بدلاً من القيم الفعلية ، لذلك أصبح إسم المقياس الإحصائي معامل إرتباط الرتب ، فإذا كان لدينا متغيــرين (y,x) مرتبان بطريقة يكون فيها معامل إرتباط الرتب محسوب بالقانون الآتي :

$$r^{1} = \frac{6\sum D^{2}}{n\left(n^{2}-1\right)}$$

عندما يكون

D = الفرق بين الرتب المرافقة لكل زوج من (y,x) .

n = عدد المشاهدات

إن قيمة (r^1) تكون (+1)(−1) يجب ملاحظة نقطتين مهمتين عند تطبيق معامل إرتباط الرتب :

أولاً – من الممكن أن يتم ترتيب المشاهدات تصاعدياً أو تنازلياً ، ولكن يجب أن تستخدم قواعد الترتيب نفسها بالنسبة للمتغيرين (y,x) .

ثانياً – اذا كان هناك متغيرين أو أكثر لهما القيمة نفسها نعطــي لهمــا معدل الرتب ، والمثال الآتي يوضح تطبيق معامل إرتباط الرتب .

مثال : الجدول الآتي يوضح كيف أن عشرة طلاب قد رتبوا إستناداً الى درجة أدائهم في العمل الصفي ، وإمتحانهم النهائي ، نريد أن نعــرف هــل أن هناك علاقة بين أداء الطلبة خلال السنة كلها وأدائهم في إمتحاناتهم النهائية :

طلبة	A	B	C	D	E	F	G	H	I	J
الترتيب على أساس العمل الصفي	**2**	**5**	**6**	**1**	**4**	**10**	**7**	**9**	**3**	**8**
الترتيب على أساس درجات الإمتحان	**1**	**6**	**4**	**2**	**3**	**7**	**8**	**10**	**5**	**9**

الحل :

الفرق بين الرتب يوضح في الجدول الآتي :

D	1	−1	2	−1	1	3	−1	−1	−2	1
D2	1	1	4	1	1	9	1	1	4	1

$$\sum D^2 = 24$$

معامل إرتباط الرتب (Rank Correlation Coefficient) هو :

$$r^1 = 1 - \frac{6\sum D^2}{n\left(n^2 - 1\right)} = 1 - \frac{6(24)}{10\left(10^2 - 1\right)} = 0.855$$

إن القيمة العالية لمعامل إرتباط الرتب (r^1) توضح أن هناك علاقة قوية أو قريبة بين الأداء في العمل الصفي والأداء في الإمتحان ، الطلبة مع أداء جيد خلال السنة في العمل الصفي سيكونون جيدين في إمتحاناتهم أيضاً والعكس بالعكس .

2-4 معاملات الإرتباط الجزئي (Partial Correlation Coefficients) .

معامل الإرتباط الجزئي يقيس العلاقة بين أي متغيرين عندما تكون بقية المتغيرات المرتبطة بهذين المتغيرين يحافظ عليها ثابتة (أي تكون ثابتة بدون تغيير) ، مثلاً لنفرض أننا نريد أن نقيس الإرتباط بين عدد كؤوس المشروبات الحارة (X_1) المستهلكة في أحد المصايف وعدد السياح (X_2) الذين يزورون ذلك المصيف ، من الواضح أن كلا المتغيرين متأثران بقوة بحالة الجو والتي

من الممكن أن نرمز لها بـ(X_3) ، على أسس مسبقة نحن نتوقع أن (X_2,X_1) مرتبطان إيجابياً : عندما يصل عدد كبير من السياح الى المصيف يجب أن نتوقع إستهلاكاً عالياً من المشروبات الحارة والعككس بالعكس ، إن حساب معامل الإرتباط البسيط بين (X_2,X_1) ربما لا يوضح لنا العلاقة الحقيقية التي تربط بين المتغيرين وذلك بسبب تأثير متغير ثالث وهو حالة الجو (X_3) ، بعبارة أخرى إن العلاقة الإيجابية أعلاه بين عدد السياح وعدد المشروبات الحارة المستهلكة من المتوقع أن تكون صحيحة إذا إفترضنا إن حالات الجو او المناخ ثابتة ، فإذا تغير المناخ ، فالعلاقة بين (X_2,X_1) ربما تنحرف الى مدى يبدو حتى علاقة سلبية ، وهكذا فإذا كان الجو حاراً ، فإن عدد السياح سيكون كبيراً ولكن بسبب الحرارة سوف يفضلون إستهلاك مشروبات باردة أكثر من المشروبات الحارة ، فإذا تجاهلنا حالة الجو ونظرنا فقط الى ($,X_1$) ، فإننا سنلاحظ إرتباطاً سالباً بين هذين المتغيرين ، وهذه العلاقة موضحة بحقيقة أن المشروبات الحارة والى المدى نفسه عدد السياح أو الزوار متأثرة بالحرارة .

من أجل أن نقيس الإرتباط الحقيقي بين المتغير(X_1) والمتغير (X_2)علينا أن نجد طريقاً يساعدنا في أن نأخذ في الحساب التغيرات الحاصلة في المتغير الثالث الذي هو (X_3)(حالات الجو) ، وهذا الطريق يمكن أن ينجز بوساطة إستخدام معامل الإرتباط الجزئي بين (X_2,X_1)عندما يكون (X_3) ثابتاً ، إن معامل الإرتباط الجزئي هذا يقرر بتعابير معاملات الإرتباط البسيط بين مختلف المتغيرات الداخلة في علاقة متعددة .

في مثالنا الحالي هناك ثلاث معاملات إرتباط بسيط :

معامل الإرتباط بين $r_{12} = X_2,X_1$

معامل الإرتباط بين $r_{13} = X_3,X_1$

معامل الإرتباط بين $r_{23}= X_3,X_2$

وهناك معاملا إرتباط جزئي :

معامل الإرتباط الجزئي بين (X_2,X_1) عندما يكون (X_3) ثابتاً = $r_{12.3}$

$$r_{12.3}=\frac{r_{12}-(r_{13})(r_{23})}{\sqrt{(1-r^2_{13})(1-r^2_{23})}}$$

معامل الإرتباط الجزئي بين (X_3,X_1) عندما يكون (X_2) ثابتاً = $r_{13.2}$

(عندما يحافظ على (X_2) ثابتاً)

$$r_{13.2}=\frac{r_{13}-(r_{12})(r_{23})}{\sqrt{(1-r^2_{12})(1-r^2_{23})}}$$

وسنعود الى مناقشة هذا الجانب من معامل الإرتباط الجزئي عند دراسة موضوع الإنحدار .

2-5 حدود نظرية الإرتباط الخطي .

إن تحليل الإرتباط له حدود أو محددات مهمة بوصفه أسلوب لدراسة العلاقات الإقتصادية (Economic Relationships) .

أولاً : إن القانون (r) يطبق فقط عندما تكون العلاقة بين المتغيرات خطية ، ولكن ربما يكون هناك متغيران مرتبطان بقوة في علاقة غير خطية ، يجب أن يكون واضحاً أن قيمة صفر لإرتباط متغيرين (Y,X) وكونهما إحصائياً مستقلان ليس شيئاً واحداً فالإرتباط يتضمن :

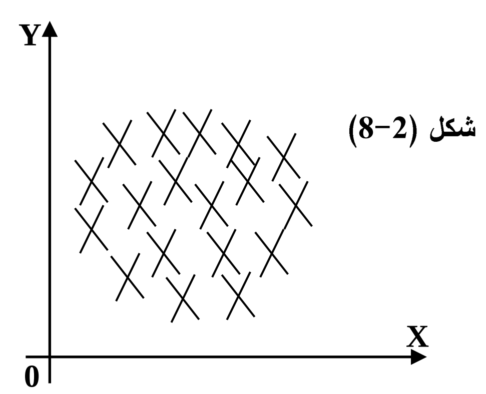
$$r=\frac{\sum xiyi}{\sqrt{\sum x^2}\sqrt{\sum y^2}}=0$$

بينما الإستقلال الإحصائي للمتغيرين (Y,X) يتضمن إحتمالية إن (Y_i,X_i) يحدثان آنياً ، أي في الوقت نفسه هي ناتج الإحتمال الفردي لكل متغير

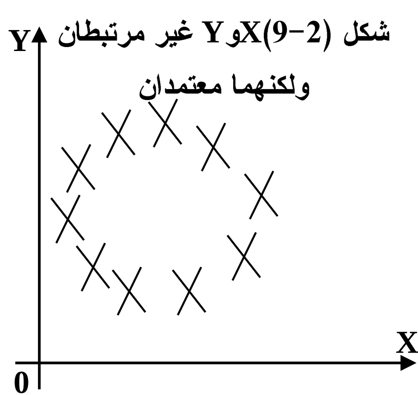
$$P (X \text{ and } Y) = P (X) . P(Y)$$

من المتغيرات لها بالتأكيد معامل تباين مشترك (Covariance) يساوي صفراً وهي بهذا غير مرتبطة ببعضها :

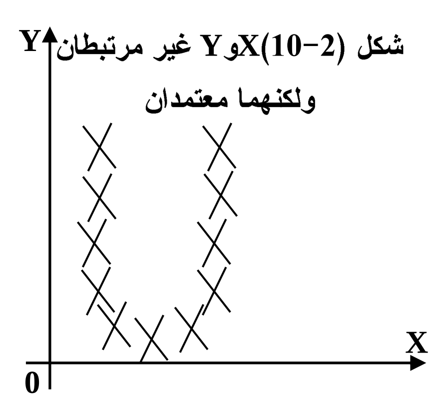
معامل الإرتباط الخطي بين متغيرين مستقلين يساوي صفراً كما في الشكل (2−8) ، ولكن معامل الإرتباط الخطي المساوي للصفر لا يتضمن بالضرورة الإستقلال (Independence) ، بكلمات أخرى فإن المتغيرات غير المرتبطة (Uncorrelated Variables) ربما تكون إحصائياً معتمدة (dependent) ، مثلاً اذا كان المتغير (X) والمتغير (Y) لهما مشاهدات أو نقاط تقع على دائرة أو على شكل حرف (U) باللغة الإنجليزية كما في الشكلين (2−9)(2−10) ، فالعلاقة بينهما تامة أو كاملة ولكن غير خطية .



شكل (2−8)



شكل (2−9)X وY غير مرتبطان ولكنهما معتمدان



شكل (2−10)X وY غير مرتبطان ولكنهما معتمدان

المتغيرات هنا إحصائياً معتمدة ، على الرغم من أن التباين المشـــترك (Covariance) ، ومعامل الإرتباط الخطي هما صفراً ، إن غيــاب الإرتبــاط الخطي (Linear Correlation) لايتضمن غياب الإعتماد بين المتغيرات .

ثانياً : إن المحدد أو المقيد الثاني لنظريــة الأرتبــاط هــو أن معامــل الإرتباط مقياس لدرجة التباين المشترك للمتغيرات لا يتضمن بالضــرورة أيــة علاقة دالية (A Functional Relationshipe) بين المتغيرات الموجودة فـــي الإرتباط ، اذ إن نظرية الإرتباط (Correlation Theory) لا تبنى أو تؤسس أو تبرهن أية علاقة سببية (A causal Relationshipe)، بين المتغيرات .

إن نظرية الإرتباط تسعى الى إكتشاف في ما إذا كان هنــاك تبــاين أو تغير مشترك موجود بين المتغيــرات ، ولكنهــا لا توضـــح أن التغيــرات أو الإنحرافات مثلاً في المتغير (Y) هي بسبب تغيرات في أو إنحرافات في المتغير (X) أو بالعكس . إن معرفة قيمة (r) لوحدها سوف لن تمكننا من التنبوء بقــيم (Y) من(X) ، إن إرتباطاً عالياً بين المتغيران (Y,X) ربما يصف لنا أي واحد من المواقف التالية :

1- التغيرات في (X)هي سبب التغيرات أو الإنحرافات في (Y).

2- التغيرات أو الإنحرافات في(Y) هي سبب التغيرات أو الإنحرافات في(X)

3- إن (X) و(Y) معتمدان معــاً (Jointly dependent) أو هنــاك إتجاهين أو طريقين للسببية (Tow-Way of Causation) وهذا يعني أن نقول أن (Y) هو سبب لــ (X) أو يتقرر بوساطة (X) ، وأيضاً (X) هي سبب لــــ (Y) أو تقرر بوساطة (Y) ، مثلاً في أي سوق من الأسواق فإن $P = F(Q)$ أو $Q = F(P)$ ولذلك نقول هناك طريقين للسببية بين الكميــة (Q) والســعر (P) بعبارة أخرى فإن (P,Q) يتقرران آنياً (Simultaneously) .

4- هناك عامل أو عنصر مشترك آخر (Z) الذي يؤثر علــى (Y,X) بطريقة توضح علاقة قريبة بينهما ، وهذا غالباً يحصل في السلاســل الزمنيــة عندما يكون هناك إرتباطاً عالياً بين (Y,X) حتى ولو كانوا مستقلين سببياً .

5- الإرتبــاط بــين (Y,X) ربمــا يكــون بســبب الصــدفة (due to chance).

وخلاصة القول أن نظرية الإرتباط لا تبني علاقة دالية ، وهذا يعني أن نظرية الإرتباط لا تعطينا أو لا تبــين أي المتغيــرات هــو المتغيــر المعتمــد (Dependent Variable) ، وأي المتغيــرات هــو المتغيــر التوضــيحي (Explanatory Variable) . من خلال البحــث الــدقيق بإســتخدام نظريــة الإقتصاد القياسي فقط نستطيع أن نصل الى نتيجة حول العلاقة الســببية بــين المتغيران (Y,X) ، وبعبارة أخرى هل أن (X) هو الســبب لــــ (Y) أم لا ، والأكثر من هذا أن تحليل الإرتباط لا يعطينا قيماً رقمية لمعلمات العلاقة ، أي لا يعطي تقديرات للميل (Slope) والتقاطع الثابت (Constant Intercept) للدالة.

إذاً يمكن أن نقول أن معامل الإرتباط الخطي يقــيس درجــة تمحــور وإستقطاب النقاط في شكل الإنتشار حول خط مستقيم ولكنه لايعطي معادلة ذلك الخط المستقيم ، وهذا يعني عدم إعطاء قيم رقميــة للمعلمــات (parameters) لتلك الدالة المعبر عنها بالخط المستقيم ، هذه المعلمات يمكن أن تكون أجزاء من المرونات أو الميل أو المضاعف ، طالما تعلق الأمر بالنظريــة الإقتصــادية ، ومعرفة القيم الرقمية لهذه المعلمات على نحو خــاص مهمــة جــداً للمنظمــين وصانعي السياسات الإقتصادية .

ومن أجل تقدير هذه المعلمات للعلاقة المعينة فإننا نطبق طرقاً مختلفة ، سوف نبدأ بدراسة طريقة المربعات الصغرى الإعتيادية للإنحدار ، وذلــك لإن هذه الطريقة هي الأبسط من جميع الطرق ، والأكثر من كله أنها تشكل الأساس لمعظم الطرق الأخرى في أساليب الإقتصاد القياسي .

الفصل الثالث

طبيعة تحليل الإنحدار

طبيعة تحليل الإنحدار

3-1 الأصل التاريخي لمصطلح الإنحدار (Regression) .

لقــد أدخــل مصــطلح الإنحــدار مــن قبــل فرانســيس كــالتون (Francis Galton) ففي بحث مشهور وجد كالتون أنه على الــرغم مــن أن الآباء والأمهات طوال القامة من المتوقع أن يكون أبنــائهم طــوالاً ، والآبــاء والأمهات قصار القامة من المتوقع أن يكون أبنائهم قصاراً ، فإن توزيع أطوال (heights) السكان لم يتغير على نحو كبير من جيل الى آخر ، والتوضيح الذي عرضه كالتون كان أن هناك ميل لمعدل طول الأطفال مع آباء وأمهات بطــول معين للحركة أو الإنحدار نحو (regress) أو بإتجاه معدل الطول للسكان كافة .

إن قانون كالتون الكوني في الإنحدار قد أثبتت صحته من قبل صــديقه كارل بيرسون (Karl Pearson) ، الذي جمع أكثر من ألف رقم حول أطــوال مجموعات من العوائل ، لقد وجد أن معدل أطوال أبناء مجموعة مــن الآبــاء طوال القامة كان أقل من طول آبائهم ، وإن معدل طول أبناء مجموعة من الآباء قصار القامة كان اطول من أطوال آباءهم ، وهكذا فإن إنحدار طــول الأبنــاء طوال وقصار القامة على معدل طول كل الرجال .

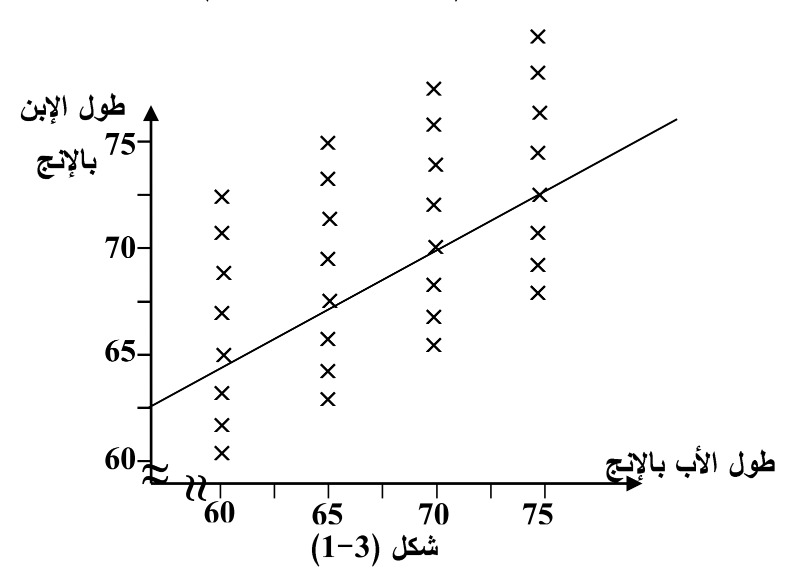
3-2 التفسير الحديث للإنحدار .

إن تحليل الإنحدار يهتم بدراسة إعتماد متغير واحد والذي هو المتغيــر المعتمد (dependent variable) على متغير واحد أو أكثر .

يسمى المتغير التوضيحي أو التفسيري (Explainatory Variable) مع رؤية أو نظرة لتقدير معدل قيمة المتغير المعتمد للمجتمع الإحصائي والتنبوء به على أساس قيم المتغير الآخر أو التوضيحي المعروفة والموجودة في عينــة إحصائية .

أمثلة :

1- إذا عدنا الى قانون كالتون في الإنحدار الكوني ، نجد كالتون كـــان يرغب في إيجاد سبب وجود الإستقرارية (Stability) فـــي توزيـــع أطـــوال السكان، ولكن في الرؤية الحديثة ليست مع هذه الصيغة في التوضيح ، ولكن مع إيجاد كيف أن معدل الطول للأبناء يتغير عندما يكون طـــول الأب معروفـــاً ، بعبارة أخرى نقول أن إهتمامنا بالتنبوء بمعدل طول الأبناء عندما تكون أطوال الآباء معروفة ، ومن أجل أن نرى كيف يمكن عمل ذلك لنأخذ الشكل البيـــاني الآتي وهو عبارة عن شكل إنتشار (Scatter Diagram) .

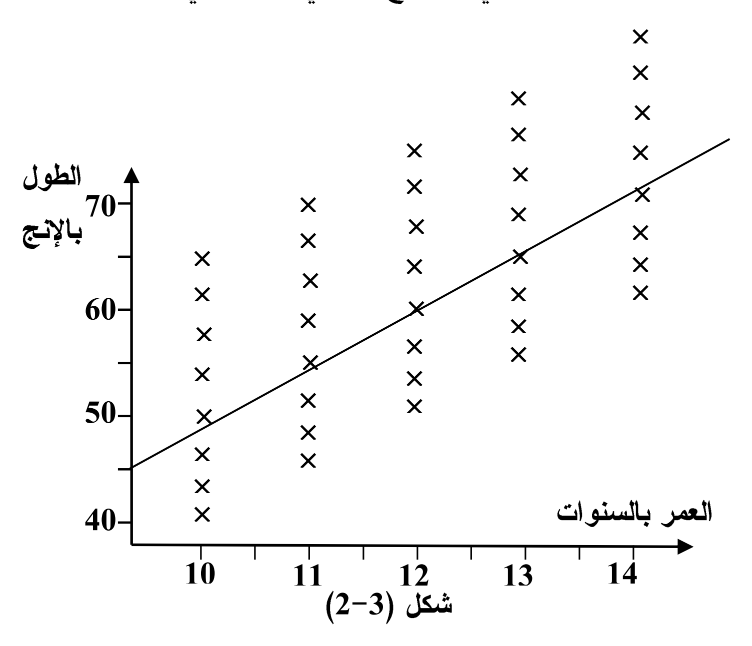


شكل (3-1)

يبين الشكل (3-1) توزيع أطوال الأبناء في مجتمع إحصائي (السكان) المقابلة لقيم معطاة لأطوال الآباء ، من الملاحظ ان القيم المقابلة لأي طول من أطوال الآباء هناك مدى (Range) أو توزيع (Distribution) لأطوال الأبناء ، ولكب يجب أن نلاحظ أن معدل طول الأبناء يزداد كلما يزداد طـــول الآبـــاء ، ولنوضح ذلك رسمنا عبر نقاط الإنتشار خطاً مستقيماً وهو يبين كيف أن معدل أطوال الأبناء يزداد مع زيادة أطوال الآباء ، وهذا الخط المستقيم كمـــا ســـوف

نرى يسمى خط الإنحدار (The Regression Line)، لاحظ أن هذا الخط لـــه ميل موجب ، ولكن الميل أقل من الواحد وهذا في تطابق مع إنحدار كالتون .

2- لنأخذ شكل الإنتشار الآتي الذي يعطينا توزيع أطــوال الأبنـــاء أو الأولاد مقاسة عند أعمار ثابتة في مجتمع سكاني إفتراضي .



شكل (3-2)

نلاحظ في الشكل (3-2) إن المقابل لأي عمر هناك مدى من الأطوال ، من الواضح أنه ليس كل الأولاد في عمر معين لهم أطوال متشـــابهة ، ولكـــن الطول على مستوى المعدل يزداد مع زيادة العمر (الى سن أو عمر معين طبعاً) وهكذا فإن معرفة العمر ربما يجعلنا قادرين على التنبوء بمعدل الطول المقابـــل أو المرافق لذلك العمر .

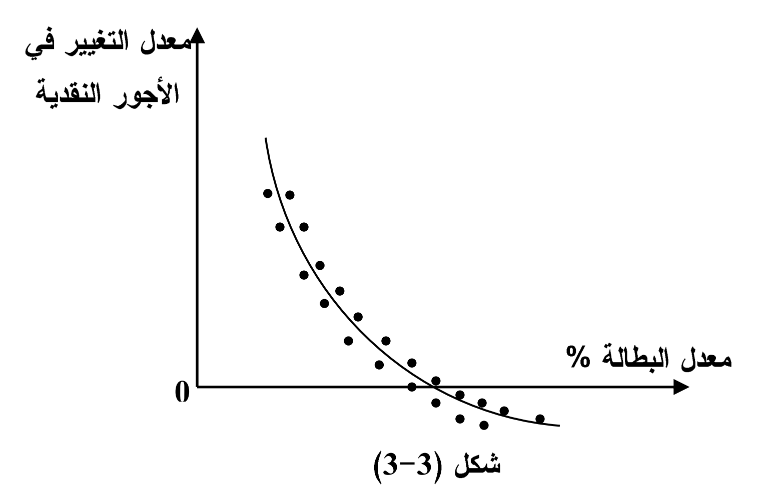
3- ننتقل الآن الى الأمثلة الإقتصادية .

ربما يكون الإقتصادي راغباً في دراســـة إعتمـــاد المصـــروف الإســـتهلاكي الشخصي على الدخل تحت التصرف بعد فرض الضريبة (الـــدخل الشخصـــي الحقيقي) إن مثل هذا التحليل ربما يكون مفيداً في تقدير الميل الحدي للإستهلاك

(Marginal Propensity to Consume) (MPC)، وذلك هو معدل التغيير في الإنفاق الإستهلاكي المرافق للتغيير في الدخل الحقيقي بقيمة دينار واحد .

إن محتكراً (A monopolist) يستطيع أن يثبت السعر أو الإنتاج (ولكن ليس كليهما) ، ربما يرغب في إيجاد إستجابة الطلب على منتوجه للتغيير في السعر ، إن مثل هذه التجربة ربما تمكنه من تقدير المرونة السعرية (الإستجابة السعرية) من قبل الطلب على المنتوج ، وربما تساعد في تقدير السعر الأكثر ربحية .

ربما يكون المختص بإقتصاد العمل يرغب أن يدرس معدل التغيير في الأجور النقدية (Money Wages) في علاقته بمعدل البطالة (Unemployment) إن البيانات التاريخية تبدو في الشكل (3−3) الآتي .



شكل (3−3)

إن المنحنى الذي يبدو في شكل الإنتشار منحنى يدعى منحنى فيلبس الذي يربط التغيرات في الأجور النقدية الى معدل البطالة ، إن مثل هذا الشكل ربما يمكننا أو يمكن المختص بإقتصاد العمل من التنبوء بمعدل التغيير في الأجور النقدية عند مستوى معين من معدل البطالة ، وإن هذه المعرفة ربما تساعد في قول شيء بخصوص العملية التضخمية في الإقتصاد الوطني ، وذلك لإن الزيادات في الأجور النقدية من المحتمل أن تعكس بزيادة في الأسعار .

3-3 الإعتماد الإحصائي مقابل الإعتماد الدالي .

من الأمثلة السابقة نلاحظ أننا في تحليل الإنحدار نهتم بما يعرف بالإعتماد الإحصائي (Statistical Dependence) وليس الإعتماد الدالي (Functional Dependence) بين المتغيرات ، ففي العلاقات الإحصائية بين المتغيرات ، فإننا بشكل رئيس نتعامل مع متغيرات عشوائية (Random) أو متغيرات إحتمالية (Stochastic Variables) وهي متغيرات لها توزيع إحتمالي ، بينما في العلاقات التامة أو التقريرية أو الدالية ، فإننا نتعامل مع متغيرات ، ولكن هذه المتغيرات ليست متغيرات عشوائية أو متغيرات إحتمالية .

إن إعتماد إنتاج المحصول على درجة حرارة الجو ومعدل سقوط الأمطار والأيام المشمسة والخصوبة مثلاً هو إعتماد إحصائي في طبيعته ، بمعنى أن المتغيرات التوضيحية على الرغم من أهميتها المؤكدة سوف لن تمكن المختص بالمحاصيل الحقلية من التنبوء بإنتاج المحصول بدقة وذلك بسبب وجود أخطاء في قياس تلك المتغيرات فضلاً عن وجود متغيرات أخرى تؤثر جماعياً على إنتاج المحصول ، ولكن ربما من الصعوبة التعرف عليها على نحو فردي .

الفصل الرابع

تحليل الإنحدار البسيط

تحليل الإنحدار البسيط

4-1 مقدمـــــــــة

نحاول في هذا الفصل مناقشة مفهوم الإنحدار بتفصيل أكثر ومن خلال هذه المناقشة ندخل الى عالم نظرية تحليل الإنحدار بأبسط صيغها وهي صـــيغة المتغيرين . والسبب في دراسة هذه الصيغة أولاً ليس بالضرورة لأنها عمليـــة كافية ، ولكن لأنها تعرض الأفكار الأساسية لتحليل الإنحدار على نحو مبســـط كلما كان ذلك ممكناً . وإن بعض هذه الأفكار يمكن أن توضح بإستخدام الأشكال البيانية ذات البعدين ، وتحليل الإنحدار المتعدد سيكون بأي شكل من الأشكال هو عبارة عن إمتداد منطقي للإنحدار البسيط .

يهتم تحليل الإنحدار على نحو كبير بتقدير أو التنبـــوء بقيمـــة الوســـط الحسابي أو معدل قيمة المتغير المعتمد على أساس قيم ثابتة ومعروفة للمتغيـــر التوضيحي أو المتغيرات التوضيحية ، ومن أجل أن نفهم كيفية عمـــل تحليـــل الإنحدار لنأخذ المثال الآتي :

لنتصور بلداً إفتراضياً يتكون من عدد سكان كلي يساوي (60) عائلة ، ونفترض أننا نرغب في دراسة العلاقة بـــين الإنفـــاق الإســـتهلاكي العـــائلي الأسبوعي (Y) والدخل تحت التصرف ، بعد فرض الضريبة الأسبوعي (X) , وعلى نحو أكثر تحديداً نفترض أننا نريد أن نتنبأ بالوسط الحســـابي لمســـتوى الإنفاق الإستهلاكي الأسبوعي العائلي عندما يكون دخل العائلة الواحدة معروفاً .

ولتحقيق الهدف نفترض أننا نقسم هذه اعوائل الـــ (60) الـــى (10) مجموعات يكون الدخل في كل مجموعة متشابهاً تقريباً بين العوائل وثم القيـــام بتفحص الإنفاق الإستهلاكي أو المصروفات الإستهلاكية عند كل مســـتوى مـــن مستويات الدخل المذكورة لمجموعات العوائل .

جدول (4−1)

X → / Y ↓	80	100	120	140	160	180	200	220	240	260
	55	65	79	80	102	110	120	135	137	150
	60	70	84	93	107	115	136	137	145	152
المصروفات	65	74	90	95	110	120	140	140	155	175
الإستهلاكية	70	80	94	103	116	130	144	144	165	178
الأسبوعية	75	85	89	108	118	135	145	145	175	180
	....	88	....	113	125	140	....	....	189	185
	....	....	....	115	....	....	....	....	....	191
المجموع	325	462	445	707	678	750	685	1043	966	1211

إن البيانات الإفتراضية في الجدول (4−1) لغرض المناقشة . إفترض إن مستويات الدخل المعطاة في الجدول (4−1) فقط مشاهدة فعلاً ، إن الجدول (4−1) يفسر كما يأتي :

عند مستوى دخل أسبوعي (80) دينار مثلاً هناك خمسة عوائل تكون مصروفاتها الإستهلاكية الأسبوعية بمدى بين (55) دينار الى (75) دينار ، وبالطريقة نفسها اذا كان مستوى الدخل (X) = (240) دينار معطى أو معروف هناك ستة عوائل تكون مصروفاتها الإستهلاكية الأسبوعية تقع بين (137) دينار الى (189) دينار .

بكلمات أخرى كل عمود في الجدول (4−1) يعطينا توزيع الإنفاق الإستهلاكي (Y) المرافق لمستوى ثابت من الدخل (X) ، بمعنى أنه يعطي التوزيع المشروط (Conditional Distribution of Y) للمتغير (Y) (المشروط على قيم معطاة من المتغير المستقل (X) الدخل) .

يجب أن نلاحظ إن البيانات في الجدول (4−1) تعبــر عــن المجتمــع الإحصائي (population) ، تستطيع أن تحسب بسهولة الإحتمالات المشروطة (Conditional probabilities) للمتغير المعتمد (Y) وهي :

$$P(Y|X)$$

إحتمالية (probability) حصول (Y) عندما تكــون (X) معروفــة أو معطاة كما يأتي :

بالنسبة الى X = 80 مثلاً هناك خمسة قيم وهي :

55 60 65 70 75

ولذلك اذا أعطينا أن X = 80 فإن إحتمالية أن نحصل على أي من القيم الخمسة المعبرة عن المصروفات الإستهلاكية هي $\frac{1}{5}$ ، وبصيغة الرموز :

$$P(Y = 55|X = 80) = \frac{1}{5}$$

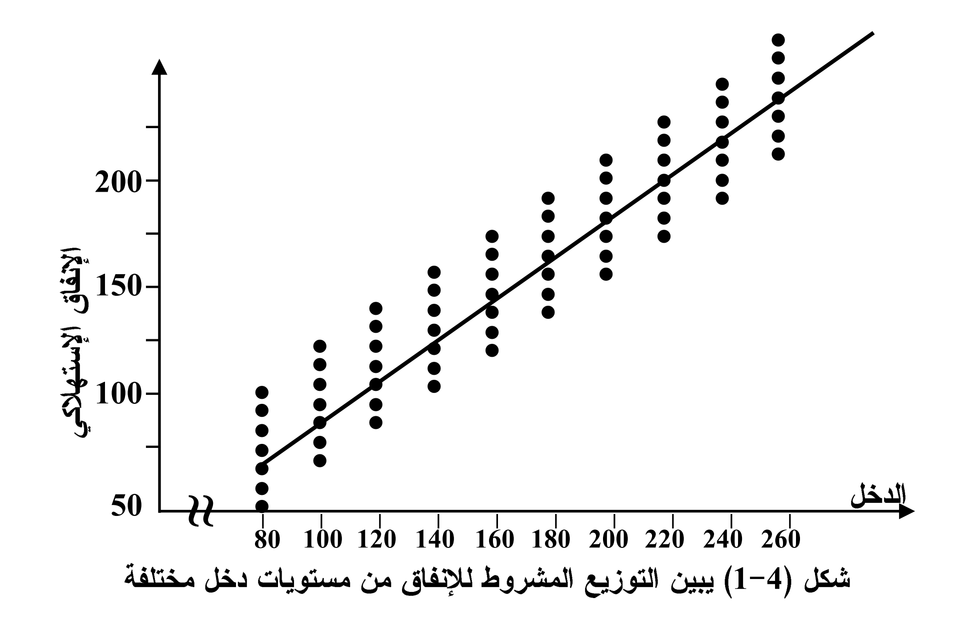
وبالطريقة نفسها فإن :

$$P(Y = 150|X = 260) = \frac{1}{7}$$

وهكذا .

إن الإحتمالات المشروطة للبيانات في الجدول (4−1) معطاة في الجدول (4−2) الآتي :

$X \rightarrow$ / $P_{\downarrow}(Y \mid X_i)$	**80**	**100**	**120**	**140**	**160**	**180**	**200**	**220**	**240**	**260**
الإحتمالات المشروطة $P(Y \mid X_i)$	$\frac{1}{5}$	$\frac{1}{6}$	$\frac{1}{5}$	$\frac{1}{7}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{5}$	$\frac{1}{7}$	$\frac{1}{6}$	$\frac{1}{7}$
	$\frac{1}{5}$	$\frac{1}{6}$	$\frac{1}{5}$	$\frac{1}{7}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{5}$	$\frac{1}{7}$	$\frac{1}{6}$	$\frac{1}{7}$
	$\frac{1}{5}$	$\frac{1}{6}$	$\frac{1}{5}$	$\frac{1}{7}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{5}$	$\frac{1}{7}$	$\frac{1}{6}$	$\frac{1}{7}$
	$\frac{1}{5}$	$\frac{1}{6}$	$\frac{1}{5}$	$\frac{1}{7}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{5}$	$\frac{1}{7}$	$\frac{1}{6}$	$\frac{1}{7}$
	$\frac{1}{5}$	$\frac{1}{6}$	$\frac{1}{5}$	$\frac{1}{7}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{5}$	$\frac{1}{7}$	$\frac{1}{6}$	$\frac{1}{7}$
	....	$\frac{1}{6}$	....	$\frac{1}{7}$	$\frac{1}{6}$	$\frac{1}{6}$	....	$\frac{1}{7}$	....	$\frac{1}{7}$
	....	....	....	$\frac{1}{7}$	....	....	....	$\frac{1}{7}$	....	$\frac{1}{7}$
	....	....	....	....	....	....	....	....	....	....
المتوسطات الحسابية لـــ(Y) المشروطة	**65**	**77**	**89**	**101**	**113**	**125**	**137**	**149**	**161**	**173**



شكل (4−1) يبين التوزيع المشروط للإنفاق من مستويات دخل مختلفة

الآن بالنسبة لكل توزيع إحتمالي مشروط لــ (Y) نستطيع أن نحسب الوسط الحسابي أو قيمة المعدل تعرف بالوسط الحسابي المشروط أو التوقع المشروط يشار اليه كما يأتي :

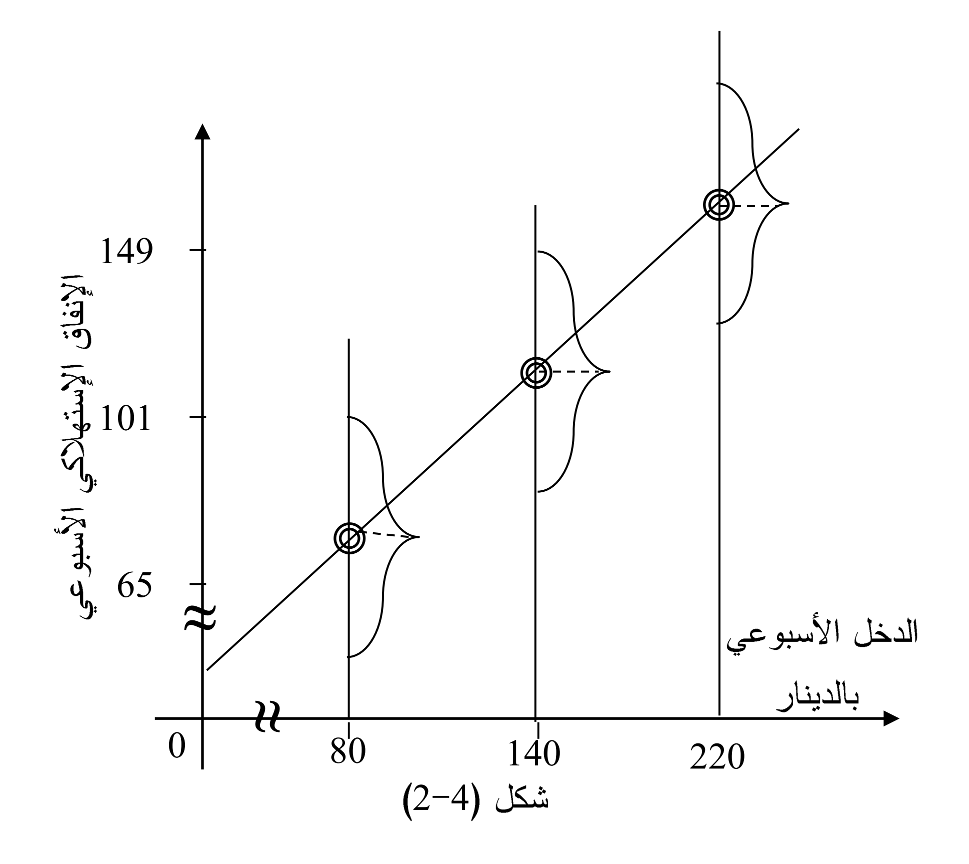
$$E\left(Y \middle| X\right)$$

ويقرأ كما يأتي : القيمة المتوقعة لــ (Y) عند مستوى معطى أو معروف من (X) ، (يجب أن نلاحظ أن القيمة المتوقعة ببساطة هي الوسط الحسابي للمجتمع الإحصائي) وبالنسبة للبيانات الإفتراضية المذكورة في المثال يمكن أن نحسب هذه التوقعات المشروطة بسهولة من خلال ضرب قيم (Y) المعطاة في الجدول (4−1) في نسب إحتمالاتها المشروطة المعطاة في الجدول (4−2) وجمع النواتج .

وللتوضيح فإن الوسط الحسابي المشروط أو توقع (Y) ، عندما تكون قيمة X = 80 هو :

$$55\left(\frac{1}{5}\right)+60\left(\frac{1}{5}\right)+65\left(\frac{1}{5}\right)+70\left(\frac{1}{5}\right)+75\left(\frac{1}{5}\right)=65$$

وهكذا فإن بقية قيم الوسط الحسابي المشروط لـــــ(Y) أو توقــع (Y) موجود في الجدول (3−2) وقبل الإستمرار في توضيح الموضوع من المفيد أن نرى بيانات الجدول (4−1) بصيغة شكل الإنتشار مبين في الشكل (4−1) ، إن شكل الإنتشار يبين التوزيع المشروط لــ(Y) المرافقة لقيم مختلفة مــن (X) ، على الرغم من وجود إنحرافات (Variations) في المصروفات الإســتهلاكية للعائلة فإن الشكل (4−1) يبين بوضوح أن الإنفاق الإستهلاكي علـــى مســـتوى المعدل (on average) يزداد كلما يزداد الدخل ، بعبارة أخرى نقول أن شـــكل الإنتشار يوضح أن قيم الوسط الحسابي المشروط لــ(Y) تزداد كلما يزداد (X) وهذا يمكن أن يبدو أكثر وضوحاً اذا ركزنا على النقاط كبيرة الحجم التي تعبر عن أوساط حسابية مشروطة مختلفة لــ(Y) ، إن شكل الإنتشـــار يبـــين هـــذه المتوسطات الحسابية المشروطة تقع تماماً على خط مستقيم مع ميل موجـــب ، وهذا الخط المستقيم يعرف أنه خط إنحدار ، وبدقة أكثر إنه منحنى إنحدار (Y) على (X) ، وهندسياً نقول أن منحنى إنحدار معين ببساطة هو عبارة عن موقع (locus) المتوسطات المشروطة أو موقع توقعات قيم المتغير المعتمد لقيم ثابتة للمتغير التوضيحي أو التفسيري ، وهذه يمكن رسمها في الشكل (4−2) الـــذي يبين أنه لكل (X) هناك قيم للمجتمع الإحصائي لــ(Y) (افترضت أنها تـــوزع توزيعاً طبيعياً) ومعها المتوسط أو الوسط الحسابي المشروط ، إن خط الإنحدار أو منحنى الإنحدار يمر عبر نقاط المتوسطات الحسابية المشروطة .



شكل (4-2)

4-2 مفهوم دالة إنحدار المجتمع الإحصائي .

ومن المناقشة السابقة وبخاصة للشكلين (4-1)(4-2) يبدو واضحاً أن كل متوسط أو وسط حسابي مشروط $E\left(Y|X_i\right)$ دالة في (X_i) وبصيغة الرمز :

$$E\left(Y|X_i\right) = f\left(X_i\right) \quad \text{.......................} \quad (4\text{-}2\text{-}1)$$

عندما تشير f(xi) الى دالة المتغير التوضيحي أو التفسيري (X_i) (في مثالنا الإفتراضي $E\left(Y|X_i\right)$ دالة خطية في (X_i) .

إن المعادلة (4-2-1) تعرف بوصفها دالة إنحدار مجتمع إحصائي بمتغيرين (population Regression Function)(PRF) ، أو إنحدار

المجتمع الإحصائي للإختصار ، وهذه العلاقة الدالية تنص على أن الوسـط أو المتوسط الحسابي للمجتمع الإحصائي لتوزيع (Y) عندما تكون (X_i) معطاة هو دالياً مرتبط بــ(X) بعبارة أخرى أنها تخبرنا كيف أن قيمة المعـدل للمجتمـع الإحصائي لــ(Y) تتغاير (Varies) مع قيم (X) .

ما هو الشكل الذي تأخذه الدالة $f(X_i)$ ؟

إن هذا السؤال مهم ، وذلك لأنه في المواقف الحقيقية نحن لا نعرف أو ليس لدينا كامل أو كل المجتمع الإحصائي جـاهزاً للإختبـار والـتفحص ، إن الشكل الدالي لدالة إنحدار المجتمع الإحصائي (PRF) هو لذلك مسألة تجريبيـة على الرغم من أنه في بعض الحالات ، ربما تخبرنا النظرية ببعض المعلومات في هذا الصدد ، مثلاً إن إقتصادياً ربما يضع الإنفاق الإستهلاكي بوصفه مرتبط خطياً بالدخل .

لذلك وكتقريب أولي أو كفرضية عملية ربما نفترض أن دالـة إنحـدار المجتمع الإحصائي (PRF-->$E(Y|X_i)$ هي دالة خطية لــ(X_i) ولنقل من نوع :

$$E(Y|X_i) = b_0 + b_1 X_i \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (2-2-4)$$

عندما تكون :

(b_0,b_1) معلمات مجهولة ولكنها معلمات ثابتة (Fixed parameters) تعرف بأنها معلمات الإنحدار ؛ كما أن (b_0) تعرف أنها معلمة التقـاطع ، وإن (b_1) تعرف أنها معامل أو معلمة الإنحـدار ، إن المعادلـة (4-2-2) ذاتهـا تعرف أنها دالة إنحدار المجتمع الإحصائي الخطية ، أو ببساطة إنحدار المجتمع الإحصائي الخطي .

4-3 التحديد الإحتمالي لدالة إنحدار المجتمع الإحصائي .

واضح من الشكل (4-1) انه كلما يزداد دخل العائلة فإن الإنفاق الإستهلاكي للعائلة على مستوى المعدل يزداد أيضاً ، ولكن ماذا حول الإنفاق الإستهلاكي لعائلة منفردة في علاقته بمستوى دخلها الثابت ؟

واضح من الجدول (4-1) والشكل (4-1) إن الإنفاق الإستهلاكي لعائلة منفردة ليس بالضرورة يزداد كلما يزداد مستوى الدخل ، مثلاً من الجدول (4-1) نلاحظ أن ما يقابل مستوى الدخل (100) دينار هناك عائلة واحدة إنفاقها الإستهلاكي (65) دينار أقل من الإنفاق الإستهلاكي لعائلتين دخلهما الأسبوعي هو (80) دينار فقط ، ولكن لاحظ أن معدل الإنفاق الإستهلاكي للعوائل مع دخل أسبوعي (100) دينار هو أكبر من معدل الإنفاق الإستهلاكي للعوائل مع دخل أسبوعي (80) دينار (65 مقابل 77) .

عندئذ ماذا يمكن أن نقول حول العلاقة بين الإنفاق الإستهلاكي العائلي (للعائلة الواحدة) المرافق لمستوى معطى من الدخل ؟ .

نشاهد من الشكل (4-1) إنه اذا اعطينا مستوى الدخل (X_i) فإن الإنفاق الإستهلاكي العائلي الفردي سوف يستقطب حول معدل الإنفاق الإستهلاكي لكل العوائل عند ذلك المستوى من الدخل (X_i) بمعنى حول التوقع المشروط ، ولذلك نستطيع أن نعرض إنحراف الإنفاق الإستهلاكي العائلي (Y_i) حول القيمة المتوقعة له كما يأتي :

$$U_i = Y_i - E(Y|X_i)$$

أو

$$Y_i = E(Y|X_i) + U_i \qquad \text{.................. } (1-3-4)$$

عندما يكون الإنحراف (U_i) متغير عشوائي غير ممكن مشاهدته يأخذ قيماً موجبة أو قيماً سالبة ، وفنياً فإن (U_i) يعرف بوصفه الإضطراب الإحتمالي (Stochastic Disturbance) أو حد الخطأ (Error Term) .

تبين لنا المعادلة (4−3−1) إن إنفاق العائلة المفــردة (عنــدما يكــون مستوى دخلها معروف) يساوي معدل الإنفاق الإستهلاكي لكل العوائل عند ذلك المستوى من الدخل زائداً بعض الكميات الموجبة أو السالبة والتي هي عشوائية، نفترض للحظات ان المتغيرات المحذوفة أو المهملة والتي تؤثر على المتغيــر المعتمد (Y) ، ولكنها غير موجودة في نموذج الإنحدار .

فإذا إفترضنا أن $E(Y|X_i)$ هو خطي في (X_i) كما في المعادلة (4−3−2) ربما نستطيع أن نكتب المعادلة (4−3−1) كما يأتي :

$$Y_i = E(Y|X_i) + U_i \;\ldots\ldots\ldots\ldots\ldots\ldots\;(4-3-2)$$
$$= b_0 + b_1 X_i + U_i$$

تؤشر أو تشير المعادلة (4−4−2) الى أن الإنفاق الإستهلاكي المشروط للعائلة الواحدة يرتبط خطياً بالدخل زائداً حد الإضطراب ، وهكذا فإن الإستهلاك للعوائل التي دخلها X = 80 دينار في الجدول (4−1) يمكن أن تعــرض كمــا يأتي :

$$\left.\begin{aligned}
Y_1 &= 55 = b_0 + b_1(80) + U_1 \\
Y_2 &= 60 = b_0 + b_1(80) + U_2 \\
Y_3 &= 65 = b_0 + b_1(80) + U_3 \\
Y_4 &= 70 = b_0 + b_1(80) + U_4 \\
Y_5 &= 75 = b_0 + b_1(80) + U_5
\end{aligned}\right\}\ldots\ldots\ldots\ldots\;(4-3-3)$$

والآن اذا أخذنا القيمة المتوقعة ل(4−3−1) على الجانبين فنحصل على ما يلي :

$$E(Y|X_i) = E[E(Y|X_i)] + E(U|X_i) \;\ldots\ldots\ldots\ldots\;(4-3-4)$$
$$= E(Y|X_i) + E(U|X_i)$$

عندما إستخدمنا حقيقة إن القيمة المتوقعة للثابت هي القيمة الثابتة نفسها يجب أن نلاحظ بإعتناء أننا في المعادلة (4−3−4) قد إخذنا الموقع المشــروط مشروطاً على مستوى معطى من (X_s) .

إن المعادلة (4-3-4) تتضمن أن :

$E(U|X_i)=0$ ....(5-3-4)

وهكذا فإن الإفتراض أن خط الإنحدار يمر عبر المتوسطات الحسابية المشروطة للمتغير المعتمد (Y) كما في الشكل (4-2) ، والذي يتضمن أن قيم المتوسطات الحسابية المشروطة للمتغير العشوائي (U_i) [مشروطة على قيمة معطاة لـ X] تساوي صفراً أو هي صفراً .

ومن المناقشة السابقة يتضح أن (4-2-2) و (4-3-2) هي أشكال متساوية اذا كانت :

$$E(U|X_i)=0$$

ولكن التحديد الإحتمالي (4-3-2) له فائدة في أنه يبين بوضوح أن هناك الى جانب الدخل متغيرات أخرى تؤثر في الإستهلاك (الإنفاق الإستهلاكي) ، وإن الإنفاق الإستهلاكي للعائلة المفردة لا يمكن أن يوضح على نحو تام وكامل من خلال المتغير أو المتغيرات الموجودة في الدالة (نموذج الإنحدار) .

4-4 دالة إنحدار العينة .

عندما اقتصر حديثنا لحد الآن على قيم المجتمع الإحصائي (population) للمتغير المعتمد (Y) المناظرة أو المرافقة الى قيم ثابتة للمتغير المستقل أو التوضيحي (X) فإننا متعمدين أن نتجنب عملية أخذ العينات (يجب أن نلاحظ ان البيانات في الجدول (4-1) تعبر عن بيانات مجتمع إحصائي وليس بيانات عينة إحصائية) ولكن الآن أصبح الوقت مهيأ لمواجهة مشكلات أخذ العينات ، وذلك لأن معظم المواقف العملية لا نملك فيها إلا قيم للمتغير المعتمد في عينة إحصائية يناظرها قيم ثابتة للمتغير المستقل أو التوضيحي أو

التفسيري (X) ، وهذا يعني أن واجبنا الآن هو تقدير دالــة إنحــدار المجتمــع الإحصائي على أساس معلومات من عينة إحصائية .

وللتوضيح لنفترض أن المجتمع الإحصائي في الجدول (4−1) كان غير معلوماً وإن المعلومات الوحيدة المتوافرة لنا كانت قد أختيرت عشوائياً ضــمن عينة إحصائية (Sample) لقيم المتغير المعتمد (Y) تناظرها قيم ثابتة للمتغيــر المستقل (X) كما هي معطاة في الجدول (4−3) الآتي :

جدول (4−3)

Y	X
70	80
65	100
90	120
95	140
110	160
115	180
120	200
140	220
155	240
150	260

ليس كما هو الوضع في الجدول (4−1) نحن الآن لدينا قيمــة واحــدة للمتغير المعتمد (Y) مقابلة لكل قيمة من المتغير المستقل (X) ، أي مقابل كــل (X) هناك قيمة واحدة (Y) .

إن الجدول (4−3) قد أختير عشوائياً من بيانات عن (Y) مناظرة نفسها من قيم المتغير المستقل أو التوضيحي (X) من المجتمع الإحصائي الموجود في الجدول (4−1) .

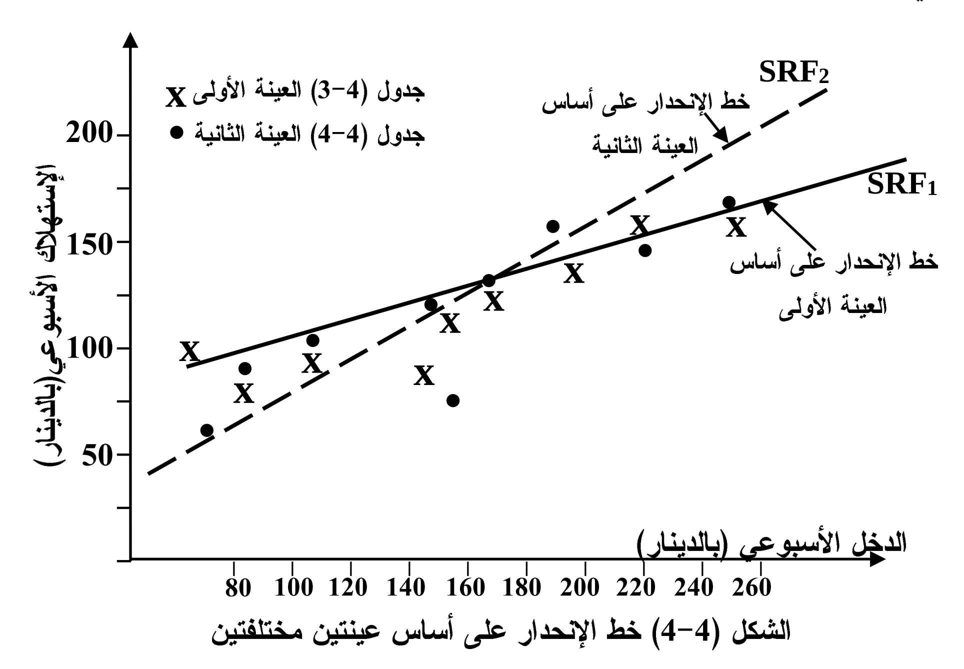
والسؤال الآن هو : من العينة الإحصائية المعبر عنها بالجـدول (4-3) هل نستطيع أن نتنبأ بمعدل الإنفاق الإستهلاكي الأسبوعي (Y) فـي المجتمـع الإحصائي بكامله المناظرة الى قيم (X) المختارة ؟ بعبارة أخرى هل نسـتطيع أن نقدر دالة إنحدار المجتمع الإحصائي (PRF) من بيانات العينة الإحصائية ؟ والجواب يمكن أن يتضمن بعض الشك في أننا ربما لا نكون قادرين على تقدير (PRF) على نحو دقيق ، وذلك بسبب التقلبات في أخذ العينات .

ومن أجل أن نرى هذا سحبنا عينة إحصـائية أخـرى مـن المجتمـع الإحصائي في الجدول (4-1) ، وكما هي معروضة في الجدول (4-4) الآتي :

جدول (4-4) عينة إحصائية عشوائية مأخوذة من بيانات جدول (4-1)

Y	X
55	80
88	100
90	120
80	140
118	160
120	180
145	200
135	220
145	240
175	260

فإذا رسمنا بيانات الجدولين (4-3)(4-4) في الشكل البياني (4-3) الآتي :



الشكل (4-4) خط الإنحدار على أساس عينتين مختلفتين

في شكل الإنتشار هناك خطي إنحدار لعينتين قد رسما فعلاً من أجل أن يوفقا منطقياً : إن دالة إنحدار العينة الأولى (SRF_1) تعتمد على العينة الأولى ودالة الإنحدار الثانية (SRF_2) تعتمد على العينة الثانية ، والسؤال هو أي من خطي الإنحدار يعبر عن خط إنحدار المجتمع الإحصائي الحقيقي ؟ .

ليس هناك طريقاً يجعلنا قادرين على نحو مطلق على تأكيد أي من الخطين المعروضين في الشكل (4-4) يعبر عن خط الإنحدار الحقيقي للمجتمع الإحصائي ، إن خطي الإنحدار في الشكل (4-4) يعرفان بوصفهما خطي إنحدار للعينات ولكنهما يمثلان في أفضل الأحوال تقريب للإنحدار الحقيقي للمجتمع الإحصائي .

إن العينة الإحصائية المقابلة للمعاملة (4-2-2) ربما تكتب كما يأتي :

$$\hat{Y}_i = \hat{b}_0 + \hat{b}_1 X_i \ldots\ldots\ldots\ldots\ldots\ldots\ldots (1-4-4)$$

عندما يكون :

$\hat{1}$ = تقرأ هات أو كاب

$\hat{Y}_i$ = مقدر لــ $E\left(Y|X_i\right)$

$\hat{b}0$ = مقدر b_0

$\hat{b}_1$ = مقدر b_1

والآن يمكننا أن نكتب المعادلة (4−4−1) بالصيغة الإحتمالية الآتية :

$$\hat{Y}_i = \hat{b}_0 + \hat{b}_1 X_i + \ell_i \ldots\ldots\ldots\ldots\ldots\ldots\ldots (1-4-4)$$

عندما يكون :

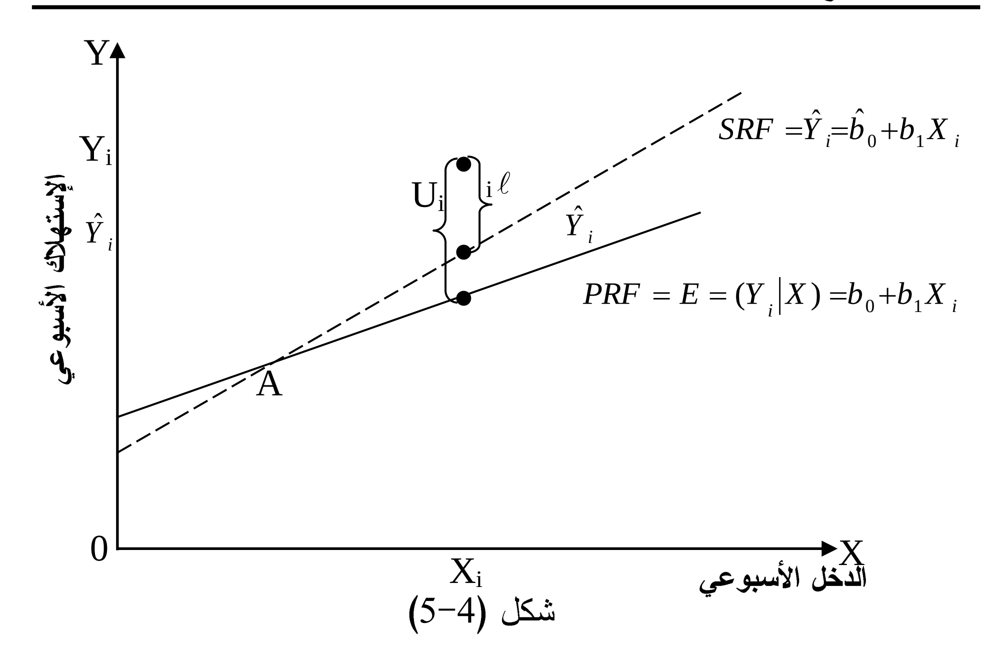
ℓ_i = يمثل حد الإضطراب أو حد البواقي (Residual Term) وفكرياً أن هذا مشابه الى (U_i) ويمكن أن يعد بوصفه تقديراً لــ(U_i) إن (ℓ_i) قد أدخــل الــى دالة إنحدار المجتمع الإحصائي .

والان لنجمع كل الذي تعلمناه ونقول أن هدفنا الأولي في تحليل الإنحدار هو لتقدير دالة أنحدار المجتمع الإحصائي (PRF) :

$$Y_i = b_0 + b_1 X_i + U_i \ldots\ldots\ldots\ldots\ldots\ldots\ldots (2-3-4)$$

على أساس دالة إنحدار العينة الإحصائية () :

$$\hat{Y}_i = \hat{b}_0 + \hat{b}_1 X_i + \ell_i \ldots\ldots\ldots\ldots\ldots\ldots\ldots (2-4-4)$$



شكل (4−5)

لكل X_i=X لدينا عينة إحصائية واحدة من المشاهدات وبصيغ دالةY_i=Y إنحدار العينة (SRF) فإن القيم المشاهدة (Y_i) يمكن أن يعبر عنها كما يأتي :

$$Y_i = \hat{Y}_i + \ell_i \quad \text{..........................(3−4−4)}$$

وبصيغة دالة إنحدار المجتمع الإحصائي (PRF) يمكن أن يعبر عنها كما يأتي :

$$\mathbf{Yi} = \mathbf{E}\left(\mathbf{Y} \middle| \mathbf{X_i}\right) + \mathbf{Ui} \quad \text{.............(4−4−4)}$$

والآن يتضح من الشكل (4−5) السابق فإن ($\hat{Y}_i$) أعطت تقديراً عالياً لقيم المتغير المعتمد الحقيقية $E\left(Y \middle| X_i\right)$ المناظرة الى قيم معطاة للمتغير المستقل أو التوضيحي (X_i) ، وبالطريقة نفسها فإن أية قيمة لـ (X_i) الى اليسار من نقطة (A) فإن دالة إنحدار العينة سوف تعطي تقديراً منخفضاً لدالة إنحدار المجتمع الإحصائي (PRF) .

إن هذه التقديرات العالية والمنخفضة هي حتمية (Invitable) بسبب التقلبات (Fluctuations) في العينات الإحصائية .

والسؤال الحاسم الآن هو : اذا قبلنا إن دالة إنحدار العينة الإحصائية (SRF) تقريب لدالة إنحدار المجتمع الإحصائي (PRF) هل يمكننا أن نبتدع قاعدة أو طريقة (Method) معينة سوف تجعل هذا التقريب أقرب ما يكون الى المجتمع الإحصائي كلما كان ذلك ممكناً ؟ .

بعبارة أخرى كيف يجب أن تبنى دالة إنحدار العينة الإحصائية ، بحيث تكون $\hat{b}_0$ قربية قدر الإمكان من المعلمة الحقيقية (b_0) وكذلك تكون ($\hat{b}_1$) قريبة قدر الإمكان من المعلمة الحقيقية (b_1) . إن الإجابة على هذا السؤال ستكون في الفصول القادمة .

الفصل الخامس

نموذج الإنحدار الخطي البسيط
وطريقة المربعات الصغرى الإعتيادية

نموذج الإنحدار الخطي البسيط وطريقة المربعات الصغرى الإعتيادية

5-1 مقدمـــــــــة .

ثمة طرق إقتصادية قياسية عدة يمكـــن بوســـاطتها أن نحصـــل علــــى تقديرات لمعلمات العلاقات الإقتصادية من المشاهدات الإحصائية ، نحاول فــــي هــــذا الفصــــل دراســــة طريقــــة المربعــــات الصــــغرى الإعتياديــــة (Ordinary Least Squares) .

إن الأسباب وراء التركيز على هذه الطريقة هي :

أولاً : إن المعلمات التي تقدر بها هذه الطريقة تكون قيم تقديراتها أكثر دقة .

ثانياً : طريقة الحساب في هذه الطريقة (OLS) على نحو عام بســيطة عند مقارنتها مع بقية الطرق الإقتصادية القياسية ، كما أنها لا تتطلب كثيراً من البيانات .

ثالثاً : إن طريقة المربعات الصغرى الإعتياديــة (OLS) كانــت قــد أستعملت في مدى واسع من العلاقات الإقتصادية مع نتائج مرضــية ، وعلــى الرغم من التحسن في أجهزة الحاسبات الإلكترونية والمعلومات الإحصائية التي تسهل إستخدام طرق قياسية أكثر تعقيداً ، نجد أن الــ(OLS) لا زالت واحــدة من الطرق الأكثر شيوعاً في الإستعمال فــي تقــدير العلاقــات فــي النمــاذج الإقتصادية القياسية .

رابعاً : إن هذه الطريقة (OLS) هي واحدة من العناصر الأساسية فــي معظم أساليب الإقتصاد القياسي .

سنبدأ هذا الفصل بنموذج الإنحدار الخطي البسيط الذي يعبر عن علاقة بين متغيرين أحدهما المتغير المعتمد والآخر المتغير التوضيحي ومرتبطان بدالة خطية .

5-2 نموذج الإنحدار الخطي البسيط .

سنوضح معنى طريقة المربعات الصغرى بمثال نظرية العرض .
إن نظرية العرض بأبسط أشكالها تفترض وجود علاقة إيجابية بين الكمية المعروضة من السلعة وسعرها مع بقاء الظروف والأشياء الأخرى ثابتة عندما يزداد السعر فإن الكمية المعروضة من السلعة ستزداد والعكس بالعكس .

بإتباع الأسلوب الإقتصادي القياسي فإن واجبنا الأول تحديد نموذج العرض وهذا يتضمن تحديد أي المتغيرات هو المتغير المعتمد ، وأي المتغيرات هو المتغير التوضيحي ، وكذلك تحديد عدد المعادلات الداخلة في النموذج والشكل الرياضي الدقيق ، وأخيراً نحدد التوقعات المسبقة التي تتعلق بإشارة وحجم المعلمات ، تزودنا النظرية الإقتصادية بالمعلومات الآتية المتعلقة بدالة العرض :

1- المتغير المعتمد هو الكميات المعروضة والمتغير التوضيحي هو السعر .

$$Y = F(X)$$

عندما يكون :

Y = الكمية المعروضة من السلعة .

X = سعر السلعة .

2- إن النظرية الإقتصادية لا تحدد كون العرض يجب أن يدرس بنموذج المعادلة الواحدة (Sigle – equation Model) ، أو بنموذج أكثر تعقيداً مكوناً من نظام من المعادلات الآنية ، ونحن هنا أخترنا أن نبدأ مع نموذج المعادلة الواحدة ، وفي مراحل متقدمة سنقوم بدراسة النماذج الأكثر تعقيداً أو النماذج الواسعة .

3- إن النظرية الإقتصادية غير واضحة حول الشكل الرياضي للنموذج (خطي أو غير خطي) لدالة العرض ، ففي علم الإقتصاد نجد أن في الكتب في

بعض الأحيان يعبر عن العرض بخط مستقيم مع ميل موجب أو منحنى مع ميل موجب ، إن الأخير يتضمن علاقة غير خطية بين الكمية والسعر ، مرة أخرى نقول أن المختص بالإقتصاد القياسي يجب أن يقوم بتقرير شكل دالة العرض ، نبدأ بإفتراض أن المتغيرات مرتبطة بأبسط شكل رياضي ممكن وهو أن العلاقة بين الكمية والسعر علاقة خطية تأخذ الشكل الآتي :

$$Y_i = b_0 + b_1X_i \ldots\ldots\ldots\ldots\ldots\ldots\ldots(1-2-5)$$

إن هذا الشكل الرياضي يتضمن طريقاً واحداً للسببية بين المتغير (Y) والمتغير (X) : السعر هو سبب التغيرات في الكمية المعروضة وليس الطريق الآخر ، إن معلمات دالة العرض هي (b_1,b_0) وهدفنا هو أن نحصل على تقديرات بقيم رقمية لهما $(\hat{b}_1,\hat{b}_0)$ بالنسبة لإشارة وحجم $(\hat{b}_0)$ نلاحظ أنها إما أن تساوي صفراً (وهذا يعني أن الكمية من السلعة المعروضة تكون صفراً عندما يكون السعر صفراً) ، أو موجباً (وهذا يعني أن كمية من السلعة تعرض حتى عندما ينخفض السعر الى الصفر) ، عادةً فإن $(\hat{b}_0)$ يجب ألا تكون بإشارة سالبة في دالة العرض .

بالنسبة لقيمة $(\hat{b}_1)$ نلاحظ أنه في حالة معينة لدالة العرض نتوقع إن إشارة $(\hat{b}_1)$ موجبة $(\hat{b}_1 > 0)$، وذلك لأن منحنى العرض بإرتفاع الميل الى الأعلى .

من المهم أن نختبر العلاقة بين المرونة السعرية للعرض والمعلمات $(\hat{b}_1)(\hat{b}_0)$ نعود الى تعريف المرونة بالتعبير الاتي :

$$Np = \frac{dQ}{dP}\cdot\frac{P}{Q}$$

$$= \frac{dY}{dX}\cdot\frac{X}{Y}$$

من دالة العرض نجد أن : $\frac{dY}{dX} = b_1$

في حساب المرونة من خط الإنحدار نستخدم $(\hat{b}_1)$ وقيم الوسط الحسابي للسعر ($\bar{X}$) والوسط الحسابي للكمية ($\bar{Y}$) في العينة .

وهكذا :
$$\hat{N}_p = \hat{b}_1 \cdot \frac{\bar{X}}{\bar{Y}}$$

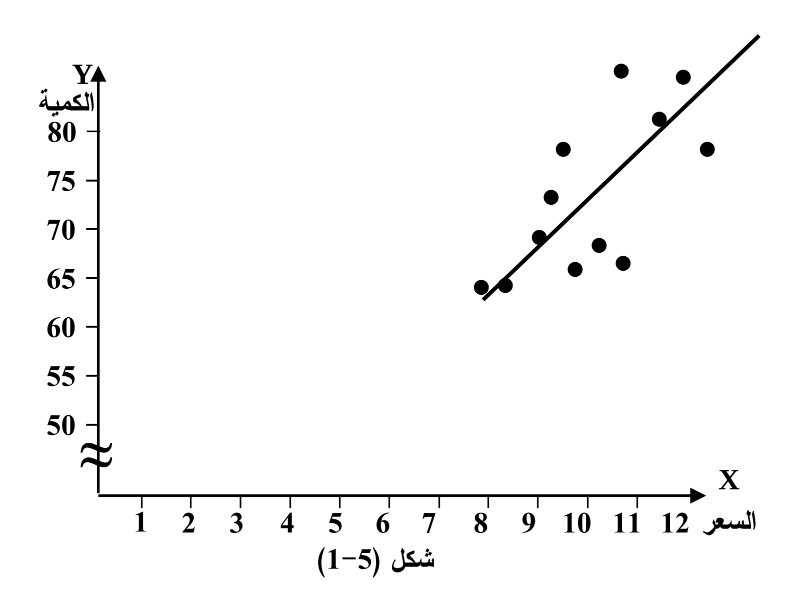
ولكن وكما سنرى لاحقاً : $\bar{Y} = \hat{b}_0 + \hat{b}_1 \bar{X}$

وهكذا عند التعويض عن ($\bar{Y}$) في معادلة المرونة نحصل على :
$$\hat{N}_p = \frac{\hat{p}_1 \bar{X}}{\hat{b}_1 \bar{X} + \hat{b}_0}$$

4- إن شكل دالة العرض المكتوب أعلاه يتضمن أن العلاقة بين الكمية والسعر هي علاقة تامة (Exact Relationship) ، وهذا يعني أن كل الإنحرافات في (Y) يعود سببها الى التغيير في(X) فقط وليس هناك عوامل أخرى تؤثر في المتغير المعتمد ، فإذا كان هذا هو الحقيقية فإن نقاط السعر والكمية اذا وضعناها في رسم بياني ستقع على خط مستقيم ، فإذا قمنا بجمع بيانات عن المشاهدات حول الكمية المعروضة فعلاً في السوق بأسعار مختلفة ، ونقوم برسم نقاط المشاهدات في شكل بياني فإننا سنلاحظ أن هذه النقاط لا تقع على خط مستقيم .

نفترض أننا لدينا عشرة نقاط عن المشاهدات حول (X) و (Y) معروضة في الجدول (5-1) الآتي :

عدد المشاهدات	الكمية (P)	السعر (X)
1	69	9
2	76	12
3	52	6
4	56	10
5	57	9
6	77	10
7	58	7
8	55	8
9	67	12
10	53	6
11	72	11
12	64	8



شكل (5−1)

شكل الإنتشار لهذه المشاهدات يبين أن العلاقة بين السعر والكمية المعروضة لها شكل خط مستقيم تقريباً كما في الشكل (5−1) ، إن إنحرافات المشاهدات أو النقاط عن الخط شبه المستقيم ربما يعود الى عدة عناصر :

1− حذف متغيرات من الدالة .

في الواقع الإقتصادي الفعلي فإن كل متغير يتأثر بعدد كبير من العناصر والعوامل ، مثلاً نموذج الإستهلاك للعائلة يتقرر بدخل العائلة والأسعار وتركيب العمر والجنس لأفراد العائلة ومستوى دخل العائلة في الماضي والأذواق والديانة والحالة التعليمية والثروة .... الخ من العوامل . ولكن لا يمكن إدخال كل المتغيرات والعناصر المؤثرة على المتغير المعتمد في الدالة وذلك لأسباب مختلفة :

أ− إن بعض العناصر ربما تكون غير معروفة حتى بالنسبة للشخص الذي هو أكثر إطلاعاً على العلاقة أو الظاهرة تحت الدراسة . إن هذا التأخر في المعرفة هو الى مدى كبير يعود الى نظرية غير مكتملة حول إنحرافات المتغيرات الإقتصادية على نحو عام .

ب- حتى عندما نعرف أن بعض المتغيرات ملائمة لإدخالها في الدالة ، ولكنن نواجه بمسألة كون بعض هذه العناصر غير قابلة للقياس إحصائياً ، هذه المتغيرات على نحو رئيس هي عناصر نفسية ، وعموماً هي عناصر أو عوامل نوعية (Qualitative Factors) مثل الأذواق والتوقعات والديانة وغيرها والتي يمكـــن التعبيـــر عنهـــا علـــى نحـــو مُرضـــي بـــالمتغيرات الصـــماء (Dummy Variables) .

ج- بعض العوامل أو العناصر هي عشوائية أو إحتمالية تبدو بطريـــق غير قابل للتنبوء ، وكذلك من حيث الزمن غير قابلة للتنبوء ، ولذلك فإن تأثيرها يمكن أن نأخذه في الحساب على نحو مُرضي ، مثلاً العـــدوى بـــالأمراض أو الهزة الأرضية وغيرها .

د- حتى ولو كانت كل العوامل أو العناصر المؤثرة معروفة ، فإن مـــا متوفر من بيانات إحصائية غالباً ما يكون غير كافٍ لقياس كل العناصر المؤثرة على العلاقة أو الظاهرة ، وهذا هو الحال على نحو خـــاص عنـــدما نســـتعمل السلاسل الزمنية والتي عادةً تكون قصيرة ، وهكذا فإنه في معظم الحالات فإن المتغيرات الثلاثة أو الأربعة المهمة تدخل على نحو واضح في الدالة .

2- السلوك العشوائي أو الإحتمالي للبشر .

إن تبعثر النقاط للمشاهدات حول خط الإنحدار المستقيم ربما يعود الـــى عنصر الخطأ الموروث في السلوك البشري إن ردود الفعل للإنسان هي الـــى مدى معين غير قابلة للتنبوء (Unpredictable) وربمـــا تســـبب إنحرافـــات (Deviations) عن نموذج السلوك "الإعتيادي" الممثل بالخط المستقيم ، مـــثلاً في لحظة نزوة ربما يغير المستهلك نموذج أو خط إستهلاكه على الـــرغم مـــن عدم تغيير الدخل أو الأسعار .

3- التحديد غير الكامل للشكل الرياضي للنموذج .

المقصود هنا أننا ربما نأخذ علاقة وندرسها على أنهـا علاقـة خطيـة ولكنها في الواقع علاقة غير خطية ، أو نحذف معادلة ونكتفي بمعادلة واحدة ، فالظواهر الإقتصادية هي أكثر تعقيداً مما يمكن أن توضـحه معادلـة واحـدة (Single Equation) ، بغض النظر عن عدد المتغيرات التوضـيحية التـي تحتويها المعادلة الواحدة ، وفي معظم الحالات فإن عدداً من المتغيرات يقـرر آنياً عن طريق نظام من المعادلات .

4- أخطاء التجميع .

غالباً ما تستعمل بيانات مجموعية او كلية مثل الإستهلاك المجموعي أو الكلي والدخل المجموعي أو الكلي ، والتي فيها نضيف حجوم للمتغيرات مـثلاً دالة الإنتاج لأي صناعة تقوم بجمع عناصر الإنتاج والإنتاج للمنظمـين غيـر المتشابهين ، إن التغيرات (changes) في توزيع الإنتاج الكلي بين الشـركات مهمة في تقرير الإنتاج الكلي .

وعلى كل حال فإن مثل هذه المتغيرات التوزيعية هي غالباً مفقودة فـي الدالة ، هناك انواع أخرى من التجميع تسبب أخطاء في العلاقة ، مثلاً التجميع على أساس المقطع العرضي (Cross - section) .... الخ .

5- أخطاء القياس .

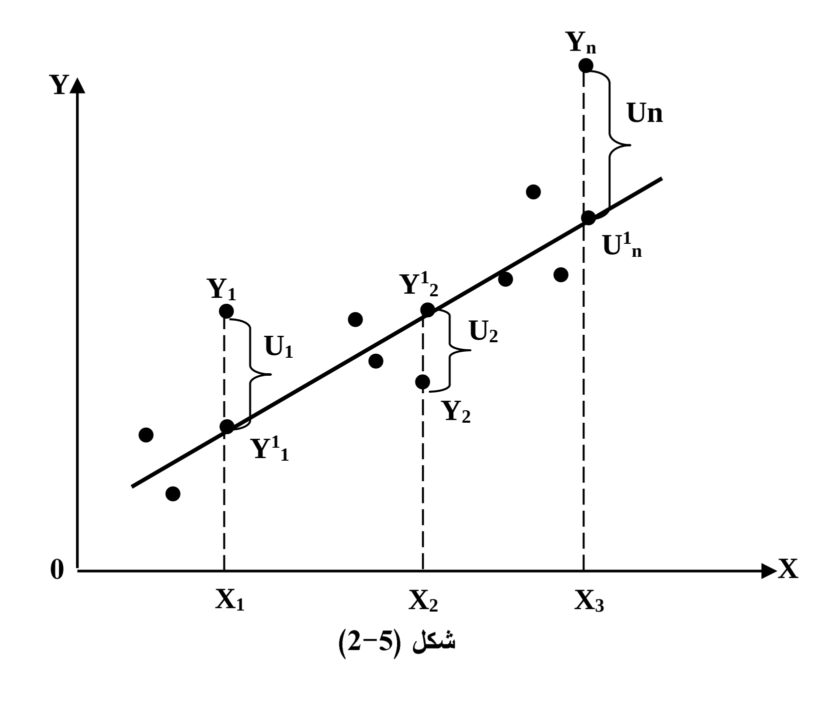
إن إنحراف نقاط المشاهدة عن خط الإنحدار المستقيم ربما يكون سـببه أخطاء في القياس هي حتمية بسبب طرق جمع البيانات ومعاملـة المعلومـات الإحصائية ، إن مصادر الأخطاء الأربعة الأولى تصور لنا شكل المعادلة خطأ ، وعـادةً يشـار لهـا (الـى هـذه الأخطـاء) أنهـا أخطـاء المعادلـة (Error of Equation) ، والمصدر الخامس للأخطاء يدعى خطأ القيـاس أو خطأ المشاهدة (Error of Measurment) ، وعادةً فإن هذين النـوعين مـن الأخطاء يكونان موجودان آنياً في المعادلة .

من أجل أن نأخذ في الحساب هذه المصادر المذكورة للأخطـــاء فإننـــا ندخل في دوال الإقتصاد القياسي المتغير العشوائي ، والذي يشار اليـــه عـــادةً بالحرف (U) ، حيث يدعى صيغة أو حـــد الخطـــأ (Error Term) ، أو حـــد الإضـــــــــطراب أو الخطـــــــــأ العشـــــــــوائي أو الإحتمـــــــــالي (Stochastic or Random Disturbance) للدالة ، بإدخال المتغير العشوائي في الدالة ، فإن النموذج يكون نموذجاً إحتمالياً بهذا الشكل :

$$Y_i = (b_0 + b_1X_i) + (U_i)\ldots\ldots\ldots\ldots\ldots\ldots(2-2-5)$$

فالعلاقة التي تربط بين المتغيرات الداخلة في العلاقة تنقسم الى جزئين أو قسمين :

الجزء الأول يعبر عنه بالخط المستقيم والجزء الثاني يعبر عنـــه عـــن طريق المتغير العشوائي (U) ، إن معنى هذين الجزئين ربما يوضح بالنظر الى الشكل (5-2) .



شكل (5-2)

إن شكل الإنتشار البياني (5−2) للمشاهدات يعبر عن العلاقة الحقيقيــة بين المتغير(Y) والمتغير (X) ، الخط المستقيم يعبر عن الجزء التام للعلاقــة ، وإنحراف المشاهدات عن الخط المستقيم تعبر عن الجزء العشوائي في العلاقة .

عندما لا يكون هناك خطأ في النموذج فإننا سنلاحظ أن النقــاط علــى الخط المستقيم (Y^1_1 و Y^1_2 و...Y^1_n) هي مرافقة لوضــع المتغيــر (X) فــي (X_1،X_2 ... X_n) ولكن بســبب وجــود المــتغــير العــشــوائي نلاحــــظ أن (Y_1 و Y_2 و .... Y_n) تــرافــق أوضــاع الــمــتــغــيــر (X) في (X1 و X2 ..... Xn) . إن هــذه النقــاط تنحــرف عــن خــط الإنحــدار (Regression Line) بالكميات (U_1,U_2,....U_n) عندما يكون (U_i) يعبر عن الخطأ المعياري الذي يرافق المتغير (Y_i) ، بعبارة أخرى فإن قيم المتغير (Y) التي تقابل أو ترافق قيم المتغير (X) ستكون على المعــدل تقــع علــى الخــط المستقيم ، ولكن كل قيمة فردية للمتغير (Y_i) سوف تنحرف عن الخط وهــذا يعتمد على قيمة (U_i) ، ولذلك فإن كل ($i = 1,2,3,......,n$) (Y_i) يمكــن أن يعبر عنها بصيغة عنصرين الأول بسبب المتغير (X) والعنصر الثاني بســبب تأثير الخطأ المعياري (U_i) .

$$Y_i = b_0 + b_1X_i + U_i$$

[الإنحراف في Y_i] = [الإنحرافات المنتظمة] + [الإنحرافات العشوائية]

أو

[الإنحراف في Y_i] = [الإنحراف الموضح] + [الإنحراف غير الموضح]

إن العنصر الأول في داخل الأقواس يعبر عــن الإنحــراف فــي (Y) موضحاً عن طريق التغيرات في (X) ، والجزء الثاني يعبر عن الإنحراف غير الواضح بأي عنصر محدد ، وهذا يعني أن الأنحراف في (Y) ســببه التــأثير العشوائي (U) .

إن معنى المتغير العشوائي (U) يبدو مرتبطاً بعبارة (مع بقاء الأشـــياء الأخرى ثابتة " Ceteris paribus ") في النظريـــة الإقتصـــادية ، فالنظريـــة الإقتصادية تفترض أن العلاقات الدالـــية بين الـــمتغيرات هي علاقـــة تـــامة (Exact Relationships) تحت ظروف بقاء الأشياء الأخرى ثابتة .

مثلاً دالة الطلب :

$$Q^d = b_0 + b_1P$$

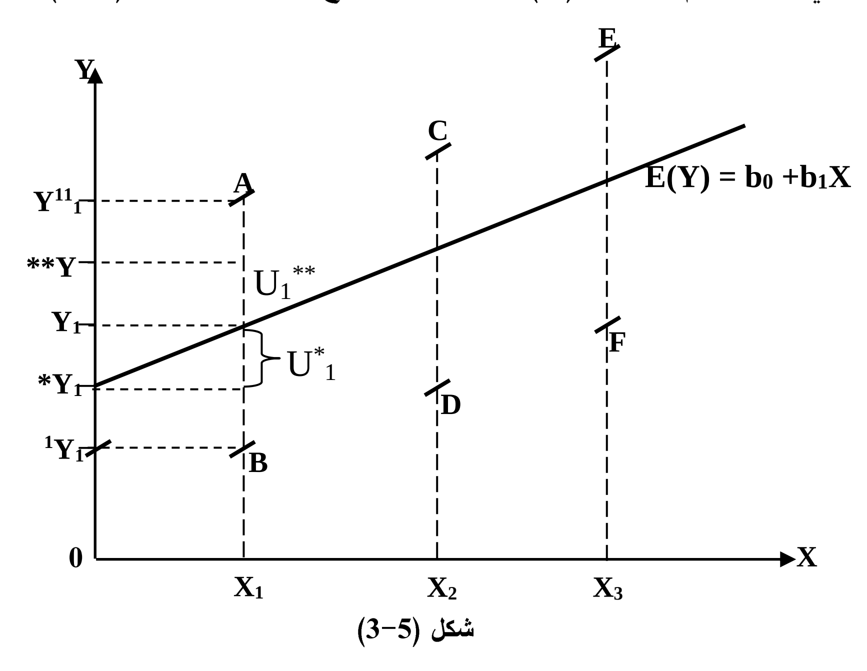
توضح بوساطة النظرية الإقتصادية ، إذ تضمنت دالة الطلب أن الكمية من سلعة معينة هي دالة خطية لسعرها فقط (بقية الأشياء الأخرى تبقـــى ثابتـــة ومتساوية) وهذا يعني أن علاقة السعرـــ الكمية تبقى صحيحة عندما تكون بقية العناصر غير الظاهرة مباشرة أو بوضوح في الدالـــة (مـــثلاً الأذواق والـــدخل وأسعار بقية السلع) تبقى بدون تغيير ، ولكن النظريات هي عبارة عن تبســـيط للعلاقات المعقدة الموجودة في العالم الحقيقي أو عالم الواقع . ولذلك فإن عبارة ((بقاء الأشياء الأخرى ثابتة)) نادراً ما تتحقق في الواقع الفعلي ، عندما نجمـــع بيانات حول الكميات المشتراة من السلعة بأسعار مختلفة ، لا نلاحظ الكميـــات التي سوف تشترى اذا كانت (بقية الأشياء الأخرى ثابتـــة) ، ولكننـــا نلاحـــظ الكميات التي تم شراؤها بينما كانت أسعار بقية السلع تتغير والـــدخل والأذواق وبقية العناصر تتغير أو في تغير مستمر . في الإقتصاد القياسي نقـــرأ العلاقـــة الحقيقية التي تربط المتغيرات بالصيغة الآتية ، إن المتغير (Y) مرتبط بالمتغير (X) بعلاقة خطية ، مع بقاء الأشياء الأخرى ثابتة .

فإذا كانت بقية العناصر أو العوامل عدا العنصر (X) بقيت ثابتة بـــدون تغيير ، فإن التغيرات في (Y) ستوضح على نحو كلي من خلال التغيرات فـــي (X) ، ولكن بقية العناصر لا تبقى ثابتة لذلك فإننا ندخل المتغير العشوائي (U) في الدالة من أجل أن نأخذ في الحساب التغييرات الحاصلة في بقية المتغيـــرات

غير الداخلة في العلاقة على نحو مباشر ، والآن يمكن ان ننظر الى الصيغة النهائية للمعادلة (5-2-2) :

$$Y_i = b_0 + b_1X_i + U_i \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(2\text{-}2\text{-}5)$$

بطريقة أخرى لكل قيمة معطاة للمتغير (Y) والمتغير (X) ربما تأخذ قيماً مختلفة بالإعتماد على القيمة المعينة للمتغير العشوائي (U)(سالباً أو موجباً) التي يأخذها هذا المتغير لكل قيمة من (X) يقابلها توزيع قيم مختلفة للمتغير العشوائي ولذلك قيم للمتغير (Y) ، إن هذا الوضع يوضحه الشكل (5-3) :



شكل (5-3)

مثلاً اذا كان سعر السلعة يساوي (X_1) فإن الكمية التي ستعرض من السلعة بهذا السعر ربما ستأخذ أية قيمة بين (${}^{1}Y_1$) و ($Y^{11}{}_1$) بالإعتماد على قيمة (U) في هذه الفترة .

فمثلاً اذا كان هناك نقص في عدد سواقي الشاحنات أو إنقطاعاً في التيار الكهربائي الذي يؤخر إستلام وتسليم السلعة (هذه الأوضاع هي أمثلة على الحوادث الحاصلة بالصدفة) ، فإن الكمية سوف لا تكون (Y_1) كما توضح ذلك

المعادلة الخطية ، ولكن كمية أصغر (Y^*_1) ، وذلك بسبب العوامل المذكورة أعلاه والتي تعطي قيمة لـ(U) هي (U^*_1) للمتغير العشوائي . من ناحية أخرى اذا كان هناك إشاعة أو توقع إن الأسعار ستنخفض للسلع البديلة أو أن منتوج جديد سيدخل السوق ، فإن المنتجين للسلعة في مثالنا سيعرضون كل الخزين منها لديهم في السوق (الكميات التي كان من الممكن عرضها في المستقبل) ، لذلك في مستوى السعر (X_1) الكمية المعروضة من السلعة ستكون (Y^{**}_1) ، وذلك بسبب التغيير في التوقعات التي تعطي قيمة للمتغير العشوائي (U) تساوي (U^{**}_1) .

من أجل تقدير المعلمات (parameters) (b_1,b_0) فإننا بحاجة الى مشاهدات حول المتغير (X) والمتغير (Y) والمتغير العشوائي (U) ، ولكن المتغير (U) لا يمكن مشاهدته مثل بقية المتغيرات التوضيحية ، وكما سنرى نستطيع الحصول على تقدير لقيم (U) بعد تقدير خط الإنحدار وحساب الإنحرافات عن خط الإنحدار ، ولذلك من أجل تقدير الدالة :

فإننا يجب أن نقدر قيم المتغير (U) ، وهذا يعني أننا يجب أن نعمل أو نضع بعض الإفتراضات (Assumptions) حول شكل توزيع المتغير(U_i)(المتوسط الحسابي والتباين والتباين المشترك) ، هذه الإفتراضات هي تقديرات للقيم الحقيقية غير المشاهدة للمتغير (U_i) .

5-3 إفتراضات نموذج الإنحدار الخطي الإحتمالي .

إن نموذج الإنحدار الخطي يعتمد على إفتراضات معينة ، وبعض هذه الإفتراضات تتعلق بتوزيع المتغير العشوائي (U) ، وبعضها الآخر حول العلاقة بين المتغير العشوائي (U) والمتغيرات التوضيحية وبعضها الآخر يتعلق بالعلاقة بين المتغيرات التوضيحية نفسها .

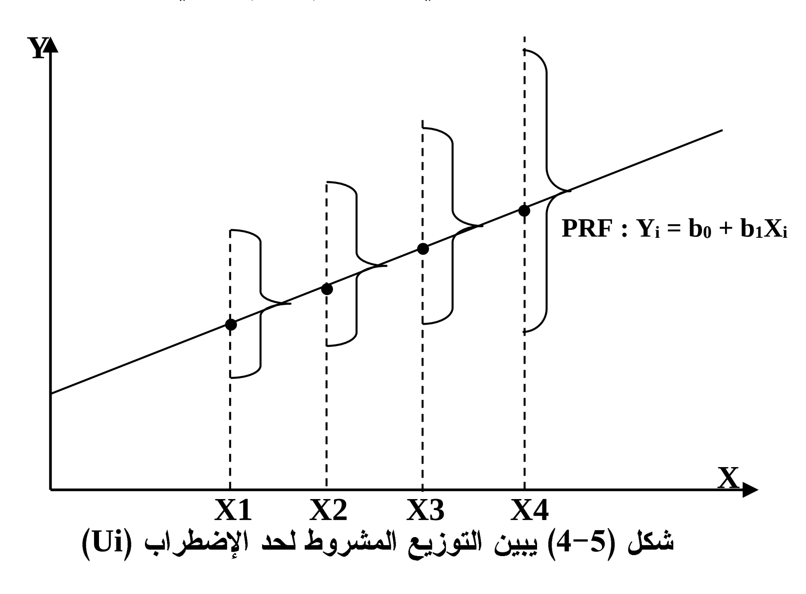
1- الإفتراض الأول :

إن (U_i) هو متغير عشوائي حقيقي ، إن القيمة التي مـــن الممكـــن أن يأخذها المتغير (U_i) في فترة واحدة تعتمد على الصدفة (chance) فربما تكون موجبة أو سالبة أو صـــفراً ، إن كـــل قيمـــة لـــــ(U_i) لهـــا نســـبة إحتمـــال (probability) في الحدوث في أي مثال ممكن .

وبصيغة المعادلة نقول :

$$E(Ui|Xi)=0..............(1-3-5)$$

وبالكلمات ينص هذا الإفتراض على أن القيمة المتوقعة المشروطة للمتغير العشوائي (U_i)(مشروطة بقيمة معطاة لـــ"X") هي صفراً ، وهندسياً فإن هذا الإفتراض يمكن أن يصور كما في الشكل (5-3) الآتي :



شكل (5-4) يبين التوزيع المشروط لحد الإضطراب (Ui)

إن الشكل (5-4) يبين عدداً قليلاً من قيم المتغير المســـتقل (X) وقـــيم المتغير المعتمد (Y) التي تناظرها من المجتمع الإحصائي للظاهرة ، وكما يبدو في الشكل ، فإن كل قيمة لـــ(Y) في المجتمع الإحصائي تناظرها قيمة معطـــاة لـــ(X) ، وتلك القيمة لـــ(Y) توزع حول الوسط الحسابي (الذي رسم بصـــيغة

دائرة داخلها نقطة فوق دالة إنحدار المجتمع الإحصائي) سوف تقع بعض قيم (Y) سوف تقع فوق قيمة الوسط أو المتوسط الحسابي وبعضها دون أو تحت قيمة الوسط الحسابي ، إن هذه المسافات للنقاط فوق وتحت قيم المتوسط الحسابي هي قيم (U_i) والذي تتطلبه المعادلة (5-2-2) هو أن قيمة الوسط الحسابي لهذه الإنحرافات (Deviations) المناظرة الى أية قيمة معطاة لـ(X) يجب أن يكون صفراً .

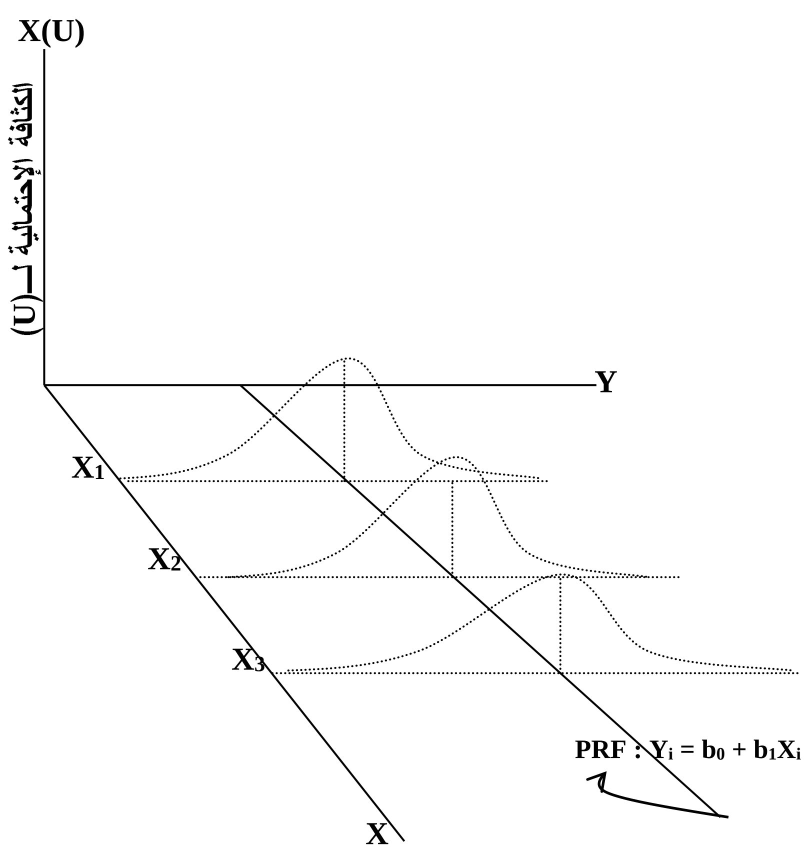
2- الإفتراض الثاني :

إن تباين (Variance) المتغير العشوائي (U_i) ثابت في كل فترة وبصيغة المعادلة نجد أن :

$$Var(U_i | X_i) = E[U_i - E(U_i)]^2$$
$$= E(U_i^2) \quad \text{بسبب الإفتراض الأول}$$
$$= \delta^2 \cdots\cdots (5-3-2)$$

عندما تكون عبارة (Var) تعني التباين (Variance) إن المعادلة (5-3-2) تنص على أن تباين المتغير (U_i) لكل (X_i) (ذلك هو التباين المشروط لـ(U_i) هو رقم موجب ثابت يساوي (δ^2) .

وفنياً فإن المعادلة (5-3-2) تعبر عن إفتراض ثبات تباين المتغير العشوائي (Homoscedasticity) ، بعبارة أخرى فإن المعادلة (5-3-2) تعني أن بيانات (Y) في المجتمع الإحصائي المناظرة لقيم مختلفة من (X) يكون لقيم (Y) التباين نفسه ، وفي الشكل البياني (5-4) يبدو هذا الوضع :

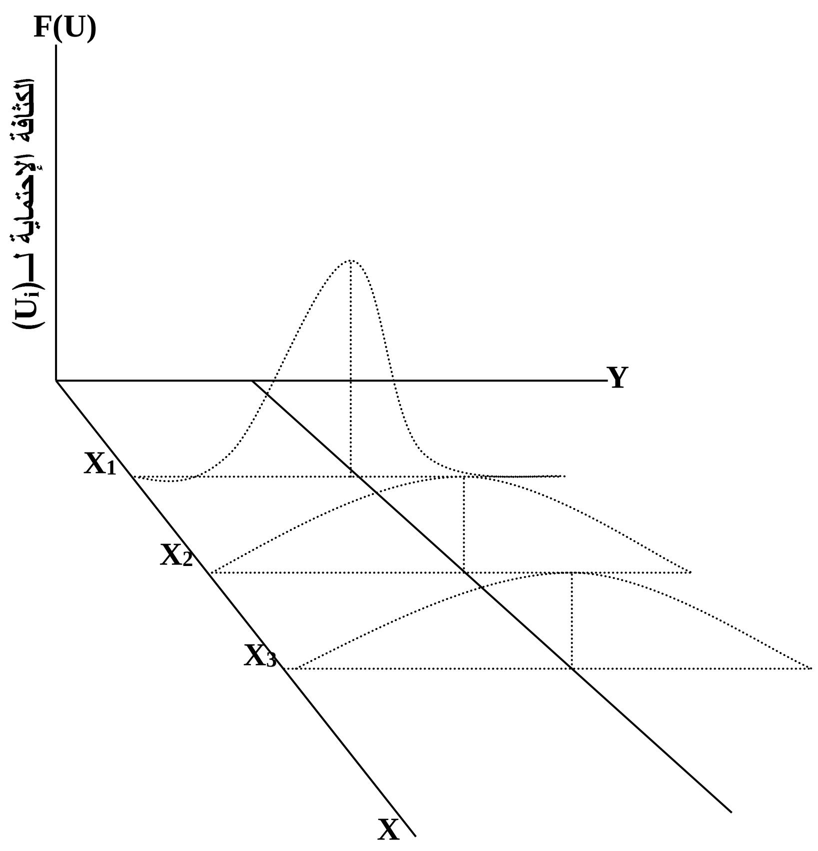


شكل (5–5) يبين ثبات تباين المتغير العشوائي

(Homoscedasicity)

وهكذا فإن تباين المتغير العشوائي (U_i) حول وسطه الحسابي ثابت في كل قيم المتغير (X) . بعبارة أخرى لكل قيم (X) فإن قيم المتغير (U) توضـــح التشتت نفسه حول وسطها الحسابي . في الشكل (5–3) ، فإن هذا الإفتـــراض يوضح الحقيقة التي تقول إن القيم التي يمكن أن يأخذها المتغير (U) تقع ضمن الحدود نفسها بغض النظر عن قيمة (X) . فعندما تكون عند قيمـــة (X_1) فـــإن المتغير العشوائي (U) يمكن ان يأخذ قيمة ضمن المدى (AB) وعند قيمة (X_2) فإن المتغير العشوائي (U) يمكن يأخذ أية قيمة ضمن المـــدى (C.D) والـــذي يساوي (AB) وهكذا .

ولكن في حالة زيادة تباين المتغير المعتمد (Y) [شـــاهدات المجتمـــع الإحصـــائي] كلمـــا يـــزداد (X) ، وهـــي حالـــة تســـمى إخـــتلاف التبـــاين (Heteroscedasticity) ولتوضيح هذه الحالة نستعين بالشكل (5−6) الآتي :



شكل (5−6) يبين إختلاف تباين المتغير العشوائي

(Hetroscedasticity)

اذا أخذنا الشكل (5−6) عندما يتغير تباين المتغير (Y) كلما يزداد (X). وهذا الوضع يسمى إختلاف أو عدم ثبات تباين المتغير العشوائي وبالمصـــطلح فإن هذا الوضع يعبر عنه بالمعادلة الآتية :

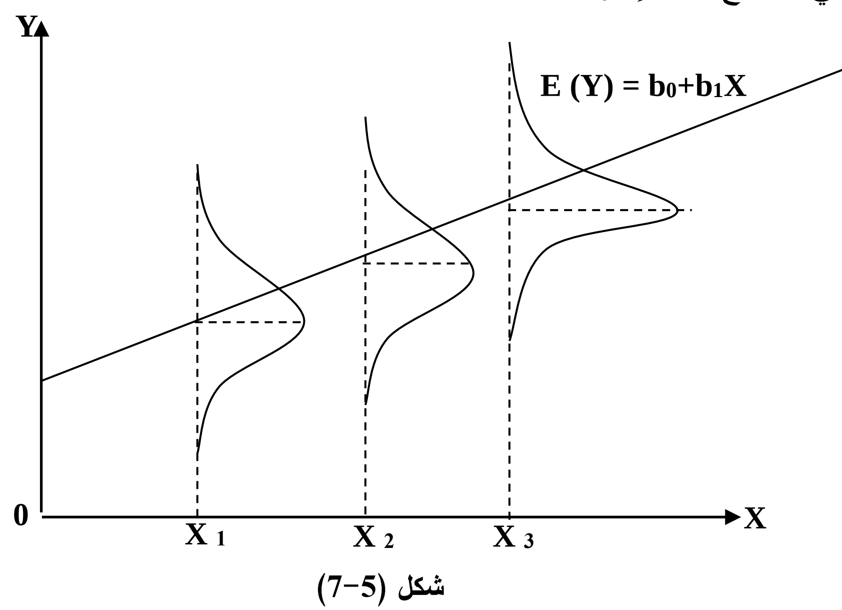
$$Var(Ui|Xi) = \delta i^{2} \cdots\cdots\cdots\cdots (3-3-5)$$

إن إضافة (i) الى (δ^2) تؤشر أن تباين (Y) للمجتمع الإحصائي لم يعد ثابتاً .

ومن أجل أن نبين الفرق بين الوضعين بوضوح ، لنجعل (Y) تعبر عن الإنفاق الإستهلاكي الأسبوعي و(X) تعبر عن الدخل الأسبوعي ، إن الشكلين (5-5) (5-6) يبينان أنه كلما يزداد الدخل فإن معدل الإنفاق الإستهلاكي يزداد أيضاً ، ولكن في الشكل (5-5) نجد أن تباين الإنفاق الإستهلاكي يبقى نفسه عند كل مستويات الدخل ، بينما في الشكل (5-6) نجد أن تباين الإنفاق الإستهلاكي يزداد مع زيادة الدخل .

3- الإفتراض الثالث .

إن المتغير العشوائي (U_i) له توزيع طبيعي (Normal Distribution) إن قيم المتغير العشوائي (لكل قيمة من X_i) لها شكل توزيع طبيعي شبيه بشكل الجرس حول وسطها الحسابي ، والشكل (5-7) الآتي يوضح هذا الإفتراض .



شكل (5-7)

إن الإفتراضات السابقة حول سلوك (توزيع) قيم المتغير العشوائي (U_i) يمكن تلخيصها بالتعبير الآتي :

$$U \sim N(O, \delta_u^2)$$

4- الإفتراض الرابع .

إن المتغير العشوائي (U) مستقل عن المتغيرات التوضيحية وهذا يعني أن المتغير العشوائي غير مرتبط بقيم المتغيرات التوضيحية ، إذ إن المتغيرالعشوائي (U) وقيم المتغيرات التوضيحية $X_1, X_2, \ldots, X_n$ لا تميل للتغير معاً أي أن التباين بينهما يساوي صفراً .

وبصيغة المعادلة فإن هذا الإفتراض يكون كما يأتي :

$$Cov(U_i, X_i) = E[U_i - E(U_i)][X_i - E(X_i)] = 0 \cdots\cdots (4-3-5)$$

ينص الإفتراض الرابع على أن حد الإضطراب (U) والمتغير التوضيحي (X) غير مرتبطان ، إن مبرر هذا الإفتراض كما يأتي :

عندما عرضنا دالة إنحدار المجتمع الإحصائي (PRF) ، كما في المعادلة (2-3-4) ، إفترضنا أن (X) و (U) [الذي ربما يعبر عن تأثير كل المتغيرات المحذوفة] لهما تأثيرين منفصلين على المتغير المعتمد (Y) . ولكن إذا كان (X) و (U) مرتبطان فليس من الممكن تقويم تأثيرهما الفردي على المتغير المعتمد (Y) .

وهكذا إذا إرتبط (X) و (U) إيجابياً فإن (X) يزداد عندما يزداد (U) وكذلك فإن (X) يتناقص عندما يتناقص (U) .

وبالطريقة نفسها إذا كان (X) و (U) مرتبطان سلبياً ، فإن (X) يزداد عندما يتناقص (U) وينحفض أو يتناقص عندما يزداد (U) . وفي كلا الحالتين من الصعوبة جداً عزل تأثير (X) وتأثير (U) على المتغير المعتمد (Y) .

إن الإفتراض الرابع هذا يتحقق تلقائيــاً اذا كــان (X) متغيــراً غيــر عشوائي او غير إحتمالي . إن نموذج الإنحدار الخطي الــذي يســتوفي هــذه الإفتراضات الأربعة السابقة يعرف بأنه نموذج الإنحدار الكلاسيكي او يــدعى نموذج الإنحدار الخطي العام . إنه نموذج كلاسيكي بمعنى أنه طور لأول مــن قبل كوز (Gauss) عام (1821) ومنذ ذلك التاريخ قد أستعمل بوصفة قاعدة او معياراً تقارن به نماذج الإنحدار التي لا تتحقق فيها الإفتراضات الأربعة لكــوز (Gaussidn).

5- الإفتراض الخامس :

إن المتغيرات التوضيحية تقاس بدون خطأ (Errors) . إن المتغير (U) يغطي تأثير المتغيرات المحذوفة ، وممكن أيضاً أن يغطي أخطاء القياس للمتغير المعتمد . وهذا يعني أننا نفترض أن المتغيرات التوضيحية حرة من الخطأ بينما قيم المتغير المعتمد لنفترض (Y) ربما تحتوي على أخطاء القياس أو ربمــا لا تحتوي على أخطاء القياس .

6- الإفتراض السادس :

إن المتغيرات التوضيحية غير مرتبطة خطياً ببعضها البعض على النحو التام .

إذا كان هناك أكثر من متغير توضيحي واحد في العلاقة ، فإننا تفترض عدم وجود علاقة تامة بين تلك المتغيرات . فإذا كانت العلاقة موجــودة بــين المتغيرات التوضيحية ، فإن هذا الوضع يدعى مشكلة الإرتباط الخطي المتعــدد (Multicollinearity) .

7- الإفتراض السابع :

إن المتغيرات الاقتصادية الكلية أو المجموعية يجب أن يكــون تجميــع بياناتها على نحو صحيح . عادة فإن المتغيرات (X) ، (Y) هي متغيرات كلية أو مجموعية تعبر عن حاصل جمع وحدات فردية . مثلاً في حالة الإستهلاك

$$C = b_0 + b_1 Y + U$$

(C) هو حاصل جمع مصروفات كل المستهلكين و (Y) هـــو حاصـــل جمع دخول الأفراد . وهكذا من المفروض أن يستخدم اسلوب تجميعي مناسب .

8- الإفتراض الثامن :

إن توزيع قيم المتغير المعتمد (Y_i) تأخـــذ شـــكل التوزيـــع الطبيعـــي (Normal Distribution) . إن شكل توزيع قيم المتغير المعتمد تقرر بوساطة شكل توزيع المتغير العشوائي (U_i) والذي هو شكل توزيـــع طبيعـــي كمـــا أن (b_1,b_0) هي معلمات ثابته لا تؤثر على توزيع قيم المتغير المعتمد (Yi) فضلاً عن أن قيم المتغيرات التوضيحية (X_i) قد وضعت قيماً ثابتة حسب الافتـــراض الخامس . ولذلك فإنها لا تؤثر على شكل توزيع قيم المتغير المعتمد (Y_i) .

5-4 : معيار المربعات الصغرى والمعادلات الاعتيادية للمربعات الصغرى (OLS)

لحد الآن قد أكملنا العمل المتضمن للمرحلة الأولى لأي بحث إقتصادي قياسي تطبيقي ويشكل أكثر وضوحاً ، فإننا قمنا بتحديد النموذج وإفتراضاته على نحو مباشر . فالخطوة اللاحقة أو التالية هي تقدير النموذج ، وهذا يعني حساب القيم الرقمية لمعلمات النموذج (Paramcters) للعلاقة الخطية :

$$Yi = b_0 + b_1 X_i + U_i$$

هذه العلاقة تكون صحيحة للمجتمع الإحصـــائية (Population) لقـــيم المتغيـــرات (X) و (Y) ، لـــذلك نســـتطيع أن نحصـــل علـــى قـــيم رقميـــة (Numerical Values) للمعلمات (b_0) ، (b_1) فقط إذا كان لدينا قيم للمتغيرات (X) ، (Y) ، (U) وهذه القيم تشكل المجتمع الإحصائي لهذه المتغيرات . ولكن بسبب أنه من غير الممكن عملياً أن نلجأ إلى الحصول على عينة (A sample) للقيم المشاهدة لـــ (X) و (Y) وتحدد توزيع المتغير العشوائي (U) ، ثم نحاول

الحصول على تقديرات للمعلمات الحقيقة في العلاقة الاقتصادية . وهذا يمكن أن نعمله عن طريق رسم خط الإنحدار من خلال مشاهدات العينـــة الإحصـــائية ، والذي نعده تقديراً أو تقريباً لخط الإنحدار الحقيقي .

إن العلاقة الحقيقة بين (X) و (Y) هي :

$$Y_i = b_0 + b_1 X_i + U_i$$

إن خط الإنحدار الحقيقي هو :

$$E(Y_i) = b_0 + b_1 X_i$$

كما أن العلاقة المقدرة (The Estimated Relationship) هي :

$$\hat{Y}_i = \hat{b}_0 + \hat{b}_1 X_i + e_i$$

وخط الانحدار المقدر (The Estimated Regression Line) هو :

$$\hat{Y}_i = \hat{b}_0 + \hat{b}_1 X_i$$

عندما يكون :

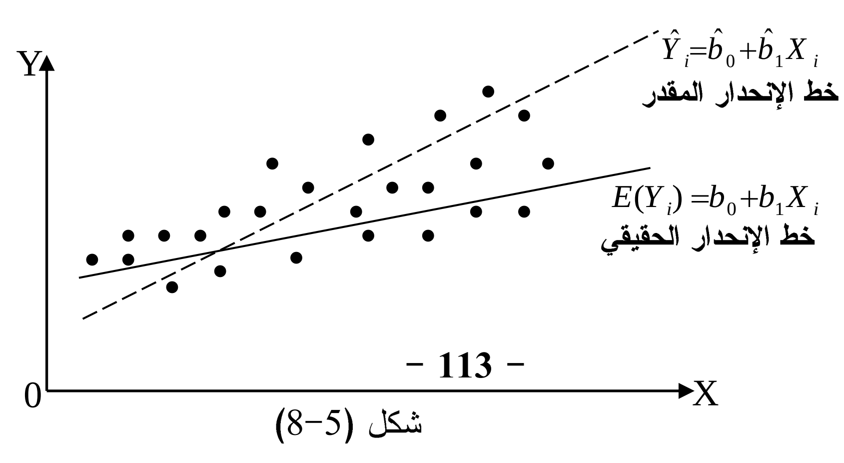
$\hat{Y}_i$= القيمة المقدرة لــ(Y) مع قيمة محددة لــ(X) .

$\hat{b}_0$= القيمة المقدرة لمعلمة التقاطع b_0 الحقيقية .

$\hat{b}_1$= القيمة المقدرة لمعلمة الميل b_1 الحقيقية .

e = القيمة المقدرة للقيمة الحقيقية للمتغير العشوائي (u) .

إن خط الإنحدار المقدر وخط الإنحدار الحقيقي يوضحهما الشكل (5-8) الآتي :



شكل (5-8)

ففي الشكل (5−8) نجد حالة العرض (Supply Function) ، ومن أجل أن نحسب القيم الرقمية للمعلمات الحقيقية (True Parameters) (b_0) و (b_1) ، يجب أن نحصل قيم للكميات المعروضة في كل مستوى من مستويات الأسعار وهذا طليعاً غير ممكن . لذلك نلجأ إلى عينة من الأسعار والكميات المشاهدة في السوق خلال فترة معينة من الزمن ونحاول أن نحصل على أفضل تقدير ممكن لدالة العرض .

إن العقبة في هذا الأسلوب هي أنه في كل عينة كعينة من الممكن أن نحصل على عدد غير محدود من خطوط الإنحدار المقدرة ، وذلك بإعطاء قيم مختلفة للمعلمات (b_0) و (b_1) .

كما لاحظنا في الفصل الرابع ، فإن دالة إنحدار المجتمع الإحصائي (PRF) هي دالة غير مشاهدة على نحو مباشر ، تقوم بتقديرها من دالة إنحدار العينة الإحصائية . ولكن السؤال المهم هو كيف يمكن أن تقرر دالة إنحدار العينة الإحصائية ؟

إذا استرجعنا نموذج الإنحدار الخطي بمتغيرين لعينة إحصائية (SRF) نستطيع أن نكتب :

$$Y_i = \hat{b}_0 + \hat{b}_1 X_i + e_i \quad \text{.........} \quad (1-4-5)$$
$$= \hat{Y}_i + e_i \quad \text{..................} \quad (2-4-5)$$

عندما تكون ($\hat{Y}_i$) قيمة مقدرة لـ(Y_i) [وسط حسابي مشروع] نستطيع أن نفرض المعادلة (5−4−2) بصيغة بديلة :

$$e_i = Y_i - \hat{Y}_i$$

$$=Y_i - \hat{b}_0 - \hat{b}_1 X_i \cdots\cdots\cdots (3-4-5)$$

وهذه المعادلة الأخيرة يتبين أن (e_i) [البواقي Residual] هي ببســاطة الفروقات بين القيم الفعلية والقيم المقدرة لــ(Y) والآن إذا أعطينا عدد (N) من أزواج من المشاهدات حول (Y) و (X) ، فإننا نرغب في تقرير دالة إنحــدار العينة (SRF) بإسلوب قدر الإمكان قريباً من القيم الفعليــة لــــ(Y) . ولهــذا الغرض ربما نتبنى المعيار الآتي :

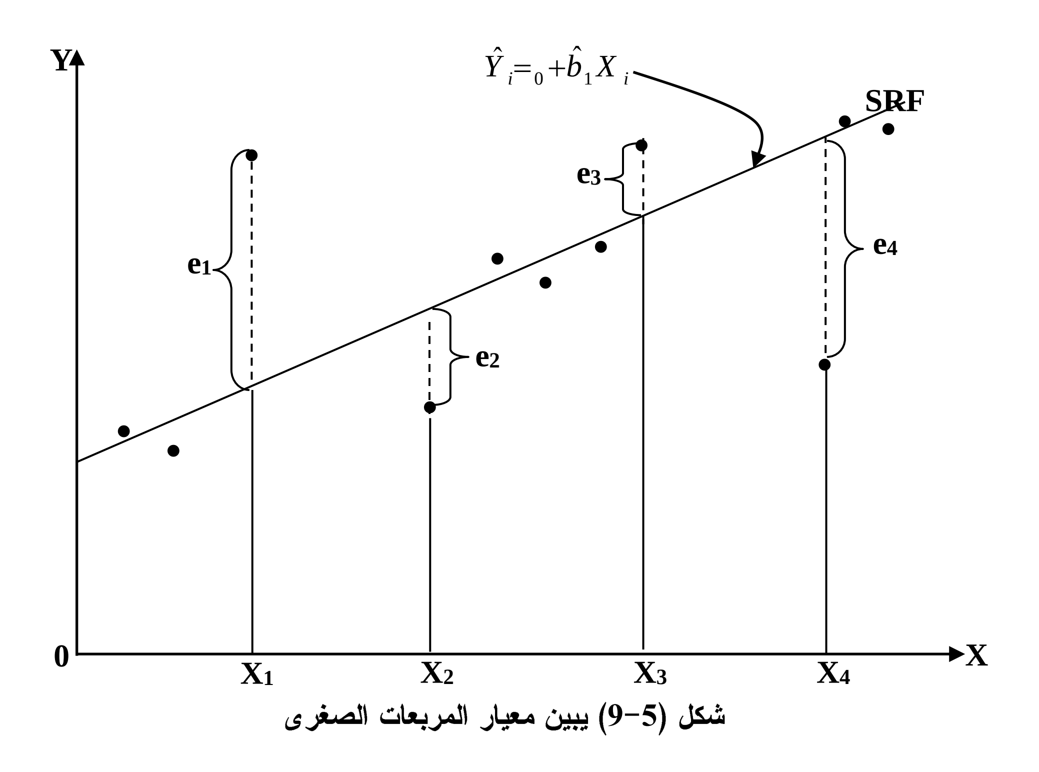
يتم إختيار دالة إنحدار العينة الإحصائية (SRF) بطريقة يكــون فيهــا مجموع البواقي :

$$\sum e_i = \sum \left(Y_i - \hat{Y} \right)$$

أصغر ما يمكن .

فإذا تبينا هذا المعيار في تصغير مجموع البواقي ($\sum e_i$) فإن الشــكل (5–4) يبين أن البواقي (e_2) و (e_3) وكذلك (e_1) و (e_2) يحصلون على الوزن نفسه في مجموع البواقي $(e_1+e_2+e_3+e_4)$ على الرغم من أن البواقي الإثنــين الأولى أكثر إقتراباً إلى دالة إنحدار العينة الإحصائية (SRF) مما هو بالنســبة للبواقي الأخيرين .بعبارة أخرى فإن كل البواقي تحصل على الأهمية المتساوية بغض النظر عن حجم إبتعاد أو إنتشار المشاهدات الفردية حول خط الإنحــدار .ونتيجة لهذا فإنه من الممكن أن يكون المجموع الجبري لــــ (e_i) صــغيراً أو صفراً . على الرغم من أن (e_i) منتشراً على نحو واسع حول دالة إنحدار العينة الإحصائية (SRF).ومن أجل أن نرى هذا لنجعل $(e_1+e_2+e_3+e_4)$ في الشــكل (5–9) تأخذ القيم (-10,+2,-2,10) .

إن المجموع الجبري لهذه البواقي هو صــفر علــى الــرغم مــن أن (e_1) و (e_4) تنتشر بإتساع كبير حول دالة إنحدار العينة الإحصائية مما إنتشــر به (e_2) و (e_3) .



شكل (5−9) يبين معيار المربعات الصغرى

نستطيع أن نتجنب هذه المشكلة إذا تبنينا معيار المربعات الصغرى (The Least – Squares Criterion) الذي ينص على أن دالة إنحدار العينة الإحصائية (SRF) يمكن أن تثبت بالطريقة الآتية :

$$\sum ei^2 = \sum (Y_i - \hat{Y}_i)^2$$
$$= \sum (Y_i - \hat{b}_0 - \hat{b}_1 X_i)^2 \cdots\cdots\cdots (4-4-5)$$

صغيراً إلى الحد الممكن . عندما يكون (e_i^2) تربيع البواقي من خلال تربيع (e_i) فإن هذه الطريقة تعطي وزناً أكبر للبواقي (e_1) و (e_4) مما تعطيه للبواقي (e_2) و (e_3) في الشكل (5−9) .

كما لاحظنا سابقاً في ظل معيار مجموع أصغر الإنحرافــات ($\sum e_i$) فإن المجموع يمكن أن يكون صغيراً حتى ولو أن (e_i) ينتشر بإتســاع حــول (SRF). ولكن هذا غير ممكــن فــي ظــل اســلوب المربعــات الصــغرى (Least-Squares) وذلك لأنه عندما يكون (e_i) كبيراً (بالقيمة المطلقــة) فــإن ($\sum ei^2$) يكون أكبر. وهناك مبرر آخر لصالح طريقة المربعات الصــغرى يكمن في حقيقة أن المعلمات المقدرة بهذه الطريقــة تمتلــك بعــض الصــفات الإحصائية المرغوبة .

من الواضح أن :

$$\sum e_i{}^2 = f(\hat{b}_0, \hat{b}_1) \cdots\cdots\cdots (5-4-5)$$

وهذه المعادلة تعني أن مجموع تربيع البواقي هو دالة للمعلمات المقدرة (b_0) و (b_1) بالنسبة لأية مجموعة بيانات معطاه تختار قيماً مختلفة لــ (b_0) و (b_1) سوف تعطي بواقي مختلفة ، ولذلك فإن قيم ($\sum ei^2$) ستكون مختلفة .

إن مبدأ المربعات الصغرى يختار $\hat{b}_0, \hat{b}_1$ بالنسبة لأي عينة بطريقــة يكون فيها مجموع تربيع البواقي ($\sum e_i^2$) صغير أو أصغر ما يمكن ، ولكــن كيف يمكن عمل ذلك ؟ إن هذا عبارة عن تمرين في حســاب التفاصــيل . إن عملية التفاصيل تنتج الصيغ الآتية لتقدير المعلمات (b_0) و (b_1) :

$$\sum Y_i = N\hat{b}_0 + \hat{b}_1 \sum X_i \cdots\cdots (6-4-5)$$

$$\sum Y_i X_i = \hat{b}_0 \sum X_i + \hat{b}_1 \sum X^2 i \cdots\cdots (7-4-5)$$

عندما يكون (N) حجم العينة (أي عدد المشاهدات) ، إن هذه المعادلات الآنية (Simultaneous Equations) تعرف بوصفها المعــادلات الطبيعيــة أوالإعتيادية (Normal Equations) .

إن هذه هذه المعلمات المقدرة ($\hat{b}_0$) و ($\hat{b}_1$) التي حصلنا عليها تعرف بمقدرات المربعات الصغرى ، وذلك لأنها أشتقت من مبدأ المربعات الصغرى ، وتتصف هذه المعلمات المقدرة بالصفات الآتية :

1. إنها تعرض فقط بصيغة كميات يمكن مشاهدتها من عينة إحصائية .
2. إنها مقدرات نقطة ، وهذا يعني إذا أعطينا عينة إحصائية ، فإن كل معلمة مقدرة سوف توفر قيمة واحدة فقط للمعلمة الملائمة للمجتمع الإحصائي .

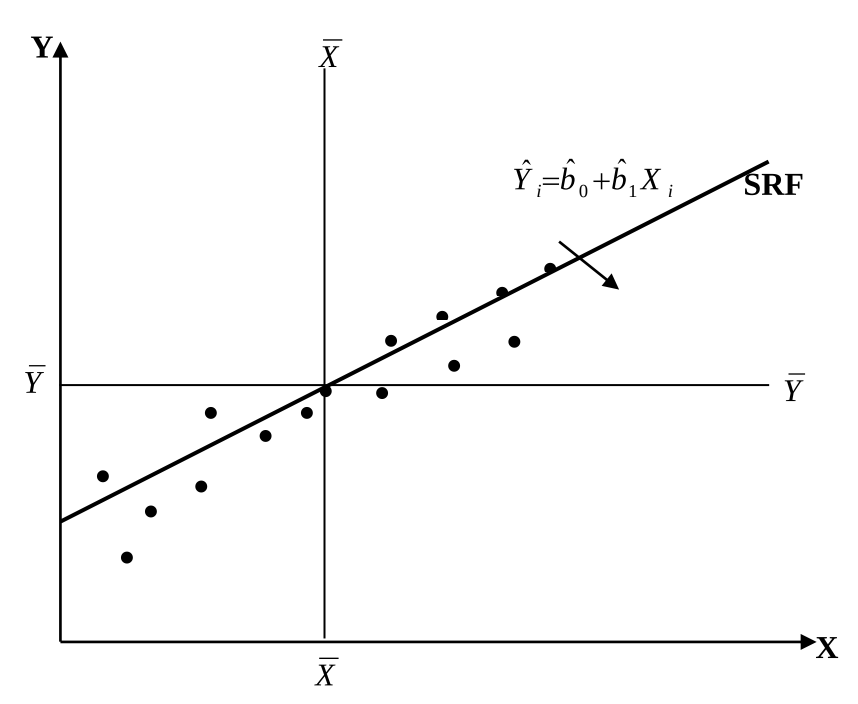
وعندما يتم الحصول على المعلمات المقدرة من البيانات المتوافرة فإن خط إنحدار العينة يمكن أن يرسم أو يوفق بسهولة . وهكذا فإن خط الإنحدار الذي حصلنا عليه يتصف بالخواص الآتية :

1. إن خط إنحدار العينة يمر عبر المتوسطات الحسابية لـ(Y) ولـ(X) .

وهذا واضح من المعادلة الآتية :

$$\overline{Y} = \hat{b}_0 + \hat{b}_1 \overline{X} \cdots\cdots (8-4-5)$$

والتي تبدو في الشكل (5–10) الآتي :



شكل (5−10) يبين أن خط إنحدار العينة الإحصائية الذي يمر عبر قيم المتوسطات الحسابية لــ(X) , (Y)

2. إن قيمة الوسط الحسابي لــ(Y) المقدرة ($\hat{Y}_i =$) تســاوي قيمــة الوسط الحسابي لــ(Y)الفعلية وذلك لأن :

$$\hat{Y}_i = \hat{b}_0 + \hat{b}_1 X_i$$
$$= (\bar{Y} - \hat{b}_1 \bar{X}) + \hat{b}_1 X_i$$
$$= \bar{Y} - \hat{b}_1 \bar{X} + \hat{b}_1 X_i$$
$$= \bar{Y} - \hat{b}_1 (X_i - \bar{X})$$
$$\hat{Y}_i = \bar{Y} - 0$$
$$\bar{\hat{Y}} = \bar{Y}$$

5-5 إشتقاق المعادلات الطبيعية

Derivation of The Normal Equations

$$\sum e_i^2 = \sum (Y_i - \hat{Y}_i)^2$$

نعوض عن قيمة ($\hat{Y}_i$) في المعادلة :

$$\sum e_i^2 = \sum \left[Y_i - (\hat{b}_0 + \hat{b}_1 X_i) \right]^2$$

$$\sum e_i^2 = \sum (Y_i - \hat{b}_0 - \hat{b}_1 X_i)^2$$

نصغر المعادلة السابقة بالنسبة الى ($\hat{b}_0$) و ($\hat{b}_1$) .

إن الشرط الضروري للتصــغير هــو أن المشــتقات الأولــى للدالــة (First Derivatives) تساوي صفراً .

$$\frac{\partial \sum e_i^2}{\partial \hat{b}_0} = 0 \qquad \frac{\partial \sum e_i^2}{\partial \hat{b}_1} = 0$$

وللحصول على المشتقات أعلاه نطبق قاعدة دالة الدالــة مــن قواعــد التفاضل وإستناداً الى هذه القاعدة :

إذا كانت : Y = F(W) وأن W = (F)X عندئذ فإن :

$$\frac{dy}{dx} = \frac{dy}{dw} \cdot \frac{dw}{dx}$$

وفي حالة الدالة السابقة لنجعل : $\left(Yi - \hat{b}_0 - \hat{b}_1 Xi \right) = w \leftarrow$

وهكذا نحصل على المشتقة الجزئية بالنسبة الى ($\hat{b}_0$) :

$$\frac{\partial \sum ei^2}{\partial \hat{b}_0} = \frac{\partial \sum (Y_i - \hat{b}_0 - \hat{b}_1 X_i)^2}{\partial \hat{b}_0} = 0$$

$$2\sum (Y_i - \hat{b}_0 - \hat{b}_1 X_i = 0) \cdot (-1) = 0$$

وبالقسمة على (2−) نحصل على المعادلة :

$$\sum(Y_i-\hat{b}_0-\hat{b}_1X_i)=0..........(1-5-5)$$

المشتقة الجزئية بالنسبة الى $(\hat{b}_1)$:

$$\frac{\partial\sum e_i^2}{\partial\hat{b}_1}=\frac{\partial\sum(Y_i-\hat{b}_0-\hat{b}_1X_i)^2}{\partial\hat{b}_1}=0$$

$$2\sum(Y_i-\hat{b}_0-\hat{b}_1X_i=0)\cdot(-Xi)=0$$

وبالقسمة على (2−) نحصل على المعادلة :

$$\sum(Y_iX_i-\hat{b}_0X_i-\hat{b}_1X_i)=0..........(2-5-5)$$

نربط المعادلتين (5−5−1)(5−5−2) ونقوم بإدخال إشارة الجمع نحصل على:

$$\sum Y_i-\sum\hat{b}_0-\sum\hat{b}_1X_i=0$$

$$\sum Y_iX_i-\sum\hat{b}_0X_i-\sum\hat{b}_1Xi^2=0$$

نطبق قواعد التجميع الإعتيادية نحصل على المعادلات الطبيعية أو الإعتيادية (Normal Equations) لطريقة المربعات الصغرى الإعتيادية (OLS) :

$$\sum Y_i=\hat{b}_0n+\hat{b}_1\sum X_i$$

$$\sum Y_iX_i=\hat{b}_0\sum X_i+\hat{b}_1\sum Xi^2$$

وبحل هاتين المعادلتين للوصول الـــى $(\hat{b}_0)$ و $(\hat{b}_1)$ نحصـــل علـــى المعلمات المقدرة لطريقة المربعات الصغرى :

$$\hat{b}_0=\frac{\sum Xi^2\sum Y_i-\sum X_i\sum X_iY_i}{N\sum Xi^2-(\sum X_i)^2}\cdots\cdots(3-5-5)$$

$$\hat{b}_1 = \frac{n\sum X_i Y_i - \sum X_i \sum Y_i}{n\sum Xi^2 - (\sum X_i)^2} \cdots\cdots (4-5-5)$$

واضح أن كل من $(\hat{b}_0)$ و $(\hat{b}_1)$ يمكن أن تقدران من خلال التعـــويض عـــن القيم $(n),(\sum X_i),(\sum Y_i),(\sum X_i Y_i),(\sum X_i^2)$ التي نحصل على قيمهـــا من مشاهدات العينة الموجودة في جدول البيانات للعينة الإحصائية .

إن القوانين المذكورة آنفاً قد عرضت بصيغ المشاهدات الأصلية للعينـــة حول كل من $(X)(Y)$ ، يمكن أن نبين أن المعلمـــات المقـــدرة$(\hat{b}_0)$ و $(\hat{b}_1)$ ربما نحصل عليها من خلال القوانين الآتية بصيغ إنحرافات المتغيـــرات عـــن وسطها الحسابي :

$$\hat{b}_0 = \overline{Y} - \hat{b}_1 \overline{X} \cdots\cdots\cdots (5-5-5)$$

$$\hat{b}_1 = \frac{\sum (X_i - \overline{X})(Y_i - \overline{Y})}{\sum (X_i - \overline{X})^2} \cdots\cdots\cdots (6-5-5)$$

البرهان

1- معلوم أن

$$\sum (X_i - \overline{X})(Y_i - \overline{Y}) = \frac{(n\sum X_i Y_i - \sum X \sum Y}{n}$$

2- ومعلوم أن

$$\sum (X_i - \overline{X})^2 = \frac{n\sum X_i^2 - (X_i)^2}{n}$$

3- نعوض في قانون ($\hat{b}_1$) نحصل :

$$\hat{b}_1=\frac{\sum(X_i-\overline{X})(Y_i-\overline{Y})}{\sum(X_i-\overline{X})^2}=\frac{\frac{n\sum X_iY_i-\sum X_i\sum Y_i}{n}}{\frac{n\sum Xi^2-(\sum X_i)^2}{n}}$$

$$\hat{b}_1=\frac{n\sum X_iY_i-\sum X_i\sum Y_i}{n\sum Xi^2-(\sum X_i)^2}$$

وللحصول على قانون ($\hat{b}_0$) نقوم بتقسيم المعادلة الطبيعية الأولى على (n) فنحصل على :

$$\frac{\sum Y_i}{n}=\frac{\hat{b}_0 n}{n}+\frac{\hat{b}_1\sum X_i}{n}$$

$$\overline{Y}=\hat{b}_0+\hat{b}_1\overline{X}$$

$$\hat{b}_0=\overline{Y}-\hat{b}_1\overline{X}$$

أو نقول إن خط الإنحدار يمر عبر نقطة تعرف من خلال المتوسـطات الحسابية للمتغيرين (Y) ، (X) .

مثال :

لتوضيح إستخدام القوانين المذكورة آنفاً سوف نقدر دالة عرض السلعة (Z) بإستخدام البيانات في الجدول (5-1) الآتي :

جدول (5−1)

N	الكمية Y_i	السعر X_i	X_i^2	X_iY_i	y_i $(Y_i-\bar{Y})$	x_i $(X_i-\bar{X})$	x_iy_i	$(X_i-\bar{X})^2$	$\hat{Y}_i=\hat{b}_0+\hat{b}_1X_i$	$e_i=(Y_i-\hat{Y}_i)$	e_i^2
1	69	9	81	621	6	0	0	0	63.00	6.00	36.00
2	76	12	144	912	13	3	39	9	72.75	3.25	10.56
3	52	6	36	312	−11	−3	33	9	52.25	−1.25	1.56
4	56	10	100	560	−7	1	−7	1	66.25	−10.25	105.06
5	57	9	81	513	−6	0	0	0	63.00	−6.00	36.00
6	77	10	100	770	14	1	14	1	66.25	10.75	115.56
7	58	7	49	406	−5	−2	10	4	56.50	1.50	2.25
8	55	8	64	440	−8	−1	8	1	59.75	−4.75	22.56
9	67	12	144	804	4	3	12	9	72.75	−5.75	30.06
10	53	6	36	318	−10	−3	30	9	53.25	−0.25	0.06
11	72	11	121	792	9	2	18	4	69.50	2.50	6.25
12	64	8	64	512	1	−1	−1	1	59.75	4.25	18.06
N=12	$\Sigma Y_i = 756$	$\Sigma X_i = 108$	$\Sigma X_i^2 = 1020$	$\Sigma X_iY_i = 6960$	$\Sigma y_i = 0$	$\Sigma x_i = 0$	$\Sigma x_iy_i = 1560$	$\Sigma x_i^2 = 48$	$\Sigma \hat{Y}_i = 756.0$	$\Sigma e_i = 0$	$\Sigma e_i^2 = 383.98$

أولاً- بإستخدام قوانين مشاهدات العينة الأصلية :

$$\hat{b}_0 = \frac{\sum X_i^2 \sum Y_i - (\sum X_i)(\sum X_i Y_i)}{n\sum X_i^2 - (\sum X_i)^2} = \frac{(1020)(756) - (108)(6960)}{(12)(1020) - (108)^2}$$

$$\hat{b}_1 = \frac{n\sum X_i \sum Y_i - \sum X_i \sum Y}{n\sum X_i^2 - (\sum X_i)^2} = \frac{(12)(6960) - (756)(108)}{(12)(1020) - (108)^2}$$

$$= \frac{1872}{576} \approx 3.25$$

ثانياً- بإستخدام قوانين إنحرافات المتغيرات عن متوسطاتها الحسابية :

$$\hat{b}_1 = \frac{\sum (X_i - \overline{X})(Yi - \overline{Y})}{\sum (Xi - \overline{X})^2} = \frac{156}{48} = 3.25$$

$$\hat{b}_0 = \overline{Y} - \hat{b}_1 \overline{X}$$

$$= 63 - (3.25)(9)$$

$$= 33.75$$

وهكذا فإن دالة العرض المقدرة هي :

$$\hat{Y}_i = 33.75 + 3.25 X_i$$

المثال الثاني : قدر دالة إنحدار الإنفاق الإستهلاكي العائلي الأســبوعي (Y) على الدخل العائلي الأسبوعي (X) من بيانات العينة الإحصائية الآتية :

Y	X
70	80
65	100
90	120
95	140
110	160
115	180
120	200
140	220
155	240
150	260

	Y_i	X_i	Y_iX_i	$X_i^{\,2}$	$\frac{x}{Xi - \overline{X}}$	$\frac{y}{Xi - \overline{Y}}$	X^2	XY_i	$\hat{Y}_i$	$\frac{e}{Yi - \overline{Y}i}$
	70	**80**	**5600**	**640**	**−90**	**−41**	**8100**	**3690**	**65.18**	**4.818**
	65	**100**	**6500**	**10000**	**−70**	**−46**	**4900**	**3220**	**75.36**	**−10.36**
	90	**120**	**10800**	**14400**	**−50**	**−21**	**2500**	**1050**	**85.54**	**4.454**
	95	**140**	**13300**	**19600**	**−30**	**−16**	**900**	**480**	**95.72**	**−0.72**
	110	**160**	**17600**	**25600**	**−10**	**−1**	**100**	**10**	**105.90**	**4.09**
	115	**180**	**20700**	**32400**	**10**	**4**	**100**	**40**	**116.09**	**−1.09**
	120	**200**	**24000**	**40000**	**30**	**9**	**900**	**270**	**126.27**	**−6.27**
	140	**220**	**30800**	**48400**	**50**	**29**	**2500**	**1450**	**136.4**	**3.54**
	155	**240**	**37200**	**57600**	**70**	**44**	**4900**	**3080**	**146.6**	**8.36**
	150	**260**	**3900**	**67600**	**90**	**39**	**8100**	**3510**	**156.8**	**−6.81**
المجموع	**1110**	**1700**	**205500**	**322000**	**0**	**0**	**33000**	**16800**	**1109.9**	**0**
الوسط الحسابي	**111**	**170**	**—**	**—**	**0**	**0**	**—**	**—**	**111**	**0**

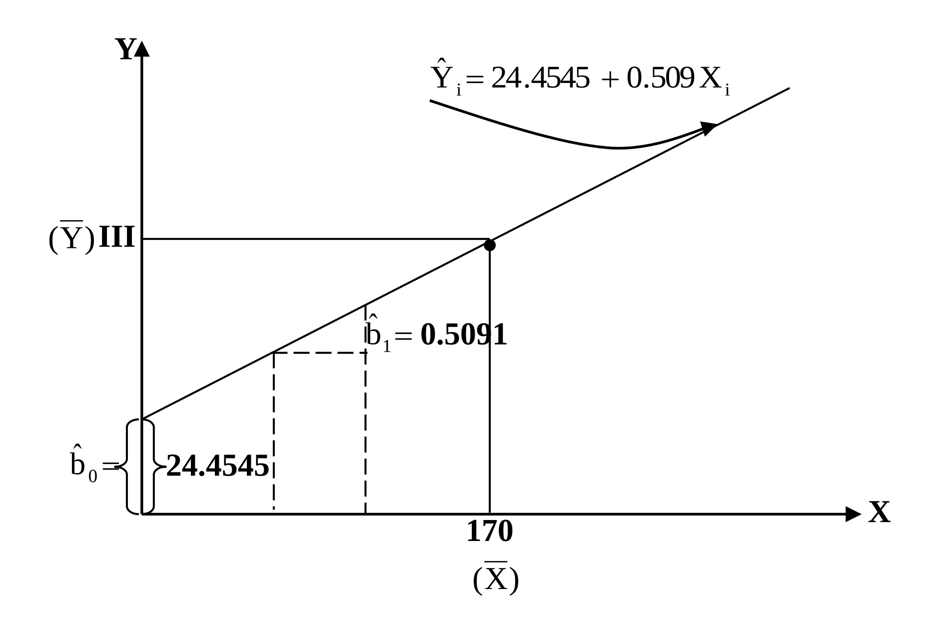
$$\hat{b}_0 = \frac{\sum(X_i - \overline{X})(Y_i - \overline{Y})}{\sum(X_i - \overline{X})} = \frac{16800}{33000}$$

$$= 0.5091$$

$$\hat{b}_0 = \overline{Y} - \hat{b}_1\overline{X}$$

$$= 111 - (0.5091)(170)$$

$$= 24.4545$$



6-5 : تقدير الدالة التي يكون تقاطعها مع المحور العمودي صفراً .

في بعض الحلات توضح لنا النظرية الإقتصادية علاقات يكون تقـاطع دالتها يساوي صفراً . وهذا يعني أن خط الإنحدار يمر من خلال نقطة إلتقـاء المحور العمودي مع المحور أو الإحداثي الأفقي ، مثلاً دالة الإنتـاج الخطيـة للمنتوجات المصنعة يجب أن يكون لها دالة مع تقاطع يساوي صـفراً ، وذلـك لأنه عندما يكون الإنتاج صفراً فإن عوامل الإنتاج أيضاً صفراً ففي هذه الحالـة يجب أن تقدر الدالة :

$$Y = b_0 + b_1 X + u$$

ولكن هنا نضع قيداً جديداً على الدالة وهو أن $0 = b_0$

لذلك فإن القانون الذي يستخدم لتقدير $\hat{b}_1$ لهذه الدالة يكون :

$$\hat{b}1 = \frac{\sum X_i Y_i}{\sum X_i^2}$$

وهذا القانون يتضمن القيم الفعلية للمتغيرات وليس الإنحرافات كما في الحالات غير المقيدة بالشرط $b_0 = 0$.

5-7 : تقدير المرونات من خط الإنحدار المقدر .

قلنا أن الدالة المقدرة :

$$\hat{Y}_i = \hat{b}_0 + \hat{b}_1 X_i$$

هي معادلة الخط الذي يكون تقاطعه مع المحور أو الأحداثي العمودي هو $\hat{b}_0$ وميله هو $\hat{b}_1$ هي عبارة عن مشتقة الدالة $\hat{Y}$ بالنسبة إلى X :

$$\hat{b}_1 = \frac{d\hat{Y}}{dX}$$

وهذه المشتقة أو الميل يبين لنا معدل التغيير في $(\hat{Y})$ الذي يصاحب التغير في (X) بكميات صغيرة جداً . يجب أن واضحاً الآن إذا كانت الدالة المقدرة هي دالة خطية للطلب على السلع أو العرض من السلع فإن المعلمة $(\hat{b}_1)$ هي ليست المرونة السعرية ، ولكن عنصر من عناصر تلك المرونة والتي تعرف بالقانون الآتي :

$$n_p = \frac{\frac{dY}{Y}}{\frac{dX}{X}} = \frac{dY}{dX} \cdot \frac{X}{Y}$$

عندما يكون : n_p = المرونة السعرية .

Y = الكمية المطلوبة أو المعروضة .

X = السعر .

يتضح أن المعلمة $\hat{b}_1$ هي العنصر $\frac{dY}{dX}$ ، لذلك نجد أو نحصل من الدالة المقدرة معدل للمرونة :

$$\eta p = \hat{b}_1 \cdot \frac{\overline{X}}{\overline{\hat{Y}}}$$
$$= \hat{b}_1 \cdot \frac{\overline{X}}{\overline{Y}}$$

عندما يكون :

$\overline{X}$ = معدل السعرات والوسط الحسابي لـــ X في العينة .

$\overline{\hat{Y}}$ = الوسط الحسابي للقيم المقدرة للمتغير المعتمد Y .

$\overline{Y}$ = الوسط الحسابي للقيم الفعلية للمتغير المعتمد Y .

يجب أن نلاحظ أن $\overline{Y} = \overline{\hat{Y}}$ ، وهذا يعني أن الوسط الحسابي للقيم المقدرة للمتغير المعتمد Y يساوي الوسط الحسابي للقيم الفعلية للمتغير المعتمد Y للعينة ، وذلك لأنه :

$$\hat{Y} = \hat{b}_0 + \hat{b}_1 X_i$$
$$\overline{\hat{Y}} = \hat{b}_0 + \hat{b}_1 \overline{X} \qquad \hat{b}_0 = \overline{Y} - \hat{b}_1 \overline{X}$$
$$= (\overline{Y} - \hat{b}_1 X) + \hat{b}_1 \overline{X}$$
$$= Y - \cancel{\hat{b}_1 X} + \cancel{\hat{b}_1 \overline{X}}$$
$$\overline{\hat{Y}} = \overline{Y}$$

في مثالنا السابق حول دالة العرض نجد أن المرونة السعرية للعرض هي :

$$n_p = (3.25)\frac{\overline{X}}{\overline{\hat{Y}}} = (3.25)\frac{9}{63}$$
$$\approx 0.46$$

الفصل السادس

الإنحدار المتعدد

الإنحدار المتعدد

6-1 : نموذج ذو متغيرين توضيحيين

Two Explanatory Variables

6-1-1 : المعادلات الطبيعية .

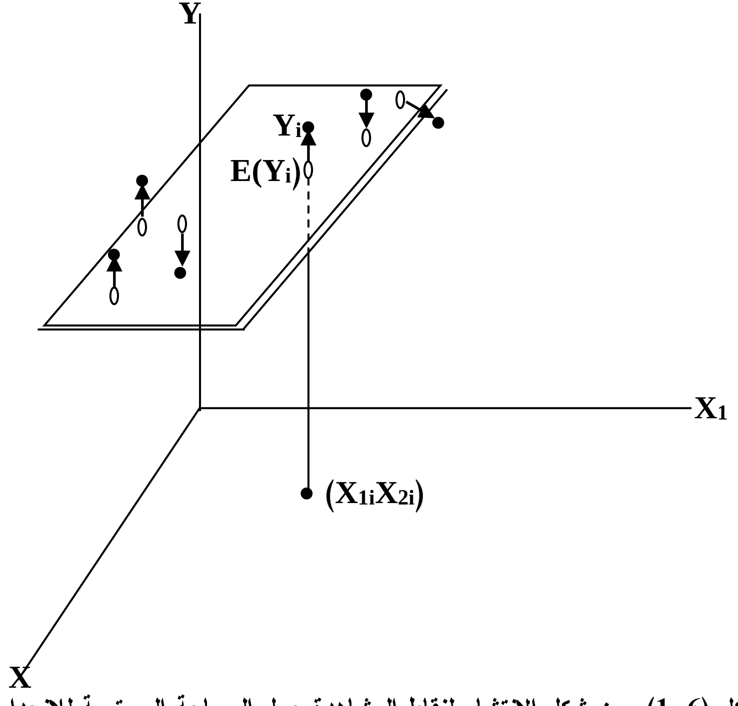
في هذه الحالة يجب أن نوضح النموذج ذو الثلاث متغيرات أحد هذه المتغيرات هو المتغير المعتمد وبقية المتغيرين هما المتغيران المستقلان . بإستخدام مثال من نظرية الطلب ، نجد أن النظرية الإقتصادية توضح لنا أن الكمية المطلوبة من سلعة معينة (Y) تعتمد على سعرها (X_1) ودخل المستهلك (X_2) . بالإمكان التعبير عن هذا بصيغة الدالة :

$$Y = f(X_1, X_2)$$

نعلم أن النظرية الإقتصادية لا تحدد لنا الشكل الرياضي لدالة الطلب لذلك نبدأ بحثنا بإفتراض إن العلاقة بين (Y) و (X_1) و (X_2) هي علاقة خطية ، ولذلك تكون لدينا دالة الطلب الآتية :

$$Y_i = b_0 + b_1 X_{1i} + b_2 X_{2i} (i = 1,2, \cdots\cdots n)$$

إن هذه العلاقة هي علاقة تامة (Exact Relationship) كما ذكرنا سابقاً ، ومعناها أن الإنحرافات في الكمية المطلوبة توضح على نحو كامل بوساطة سعر السلعة ودخل المستهلك فقط . فإذا كان هذا الشكل للعلاقة هو صحيح فإن أية مشاهدة حول (Y) و (X_1) (X_2) تعطينا نقطة تقع على المساحة المستوية (aplane) للمتغيرات (Y) و (X_1) و (X_2) .



شكل (6−1) يبين شكل الإنتشار لنقاط المشاهدة حول المساحة المستوية للإنحدار
(The Regression Plane)

فإذا جمعنا بيانات عن المشاهدت لهذه المتغيرات خلال مدة زمنية معينة ونقوم برسم نقاط المشاهدات بشكل الإنتشار ، فسوف نلاحظ أن بعض النقاط لا تقع على المساحة المستوية (plane) ، وإن بعضها يقع على المساحة المستوية وإن البعض الآخر سيقع إما فوق المساحة المستوية أو تحتها . إن هذا الإنتشار بسبب عدد من العوامل أو العناصر المحذوفة من الدالة ، فضلاً عن أنواعاً أخرى من الخطأ التي تم شرحها في الفصول السابقة .

إن تأثير مثل هذه العوامل المحذوفة يمكن أخذه في الحساب بوساطة إدخال المتغير العشوائي (u) (Random Variable) في الدالة والتي ستصبح دالة إحتمالية (Stochastic Function) .

$$\hat{Y}_i = (\hat{b}_0 + \hat{b}_1 X_{1i} + \hat{b}_2 X_{2i}) + (u_i)$$

العنصر النظامي العنصر العشوائي

على أسس مسبقة نتوقع أن المعلمة ($\hat{b}_1$) سيكون لها إشـــارة ســـالبة ، وكما هو معروف من قانون الطلب ، بينما ($\hat{b}_2$) من المتوقع أن يكـــون لهـــذه المعلمة إشارة موجبة ، وذلك لأن الكمية المطلوبة من السلع الإعتياديـــة تتغيـــر بالإتجاه نفسه الذي يتغير به الدخل .

إن الإفتراضات بالنسبة للإنحدار المتعدد هي الإفتراضات نفسها بالنسبة للإنحدار البسيط الذي مر ذكره في الفصل السابق .

بعد تحديد النموذج نقوم بإستخدام عينة إحصائية متكونة من مشـــاهدات حول (Y) و (X_1) و (X_2) نحصل على تقديرات للمعلمـــات الحقيقيـــة : b_0 و b_1 و b_2 .

$$\hat{Y}_i = \hat{b}_0 + \hat{b}_1 X_{1i} + \hat{b}_2 X_{2i}$$

عندما تكون :

$\hat{b}_0$ و $\hat{b}_1$ و $\hat{b}_2$ هي تقديرات للمعلمات الحقيقية b_0 و b_1 و b_2 للعلاقة .

وكما مر في الإنحدار البسيط ، فإن التقديرات يمكن الحصول عليهـــا بواســـطة تصغير (Minimising) مربع الإنحرافات أو مربع البواقي .

$$\sum_{i=1}^{n} \ell^2 = \sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2 = \sum_{i=1}^{n} (Y_i - \hat{b}_0 - \hat{b}_1 X_{1i} - \hat{b}_2 X_{2i})^2$$

من أجل الحصول على المعادلات الطبيعية الـــثلاث يجـــب أن نجـــري عملية التفاضل الجزئي للدالة بالنسبة إلى $\hat{b}_0$ و $\hat{b}_1$ و $\hat{b}_2$:

أولاً – المشتقة الجزئية لـــ $\hat{b}_0$:

$$\frac{\partial \sum \ell_i^2}{\partial \hat{b}_0} = \frac{\partial \sum (Y_i - \hat{b}_0 - \hat{b}_1 X_{1i} - \hat{b}_2 X_{2i})^2}{\partial \hat{b}_0} = 0$$

$$= -2\sum (Y_i - \hat{b}_0 - \hat{b}_1 X_{1i} - \hat{b}_2 X_{2i}) = 0$$

$$= \sum Y_i - \sum \hat{b}_0 - \sum \hat{b}_1 X_{1i} - \sum \hat{b}_2 X_{2i} = 0$$

$$\sum Y_i = n\hat{b}_0 + \hat{b}_1 \sum X_{1i} + \hat{b}_2 \sum X_{2i} \quad \ldots\ldots\ldots\ldots\ldots (1)$$

ثانياً – المشتقة الجزئية لـــ $\hat{b}_1$:

$$\frac{\partial \sum \ell_i^2}{\partial \hat{b}_1} = \frac{\partial \sum (Y_i - \hat{b}_0 - \hat{b}_1 X_{1i} - \hat{b}_2 X_{2i})^2}{\partial \hat{b}_1} = 0$$

$$= -2\sum X_{1i}(Y_i - \hat{b}_0 - \hat{b}_1 X_{1i} - \hat{b}_2 X_{2i}) = 0$$

$$\sum X_{1i} Y_i - \sum \hat{b}_0 X_{1i} - \sum \hat{b}_1 X_{1i}^2 - \sum \hat{b}_2 X_{1i} X_{2i} = 0$$

$$\sum X_{1i} Y_i = \hat{b}_0 \sum X_{1i} + \hat{b}_1 \sum X^2{}_{1i} + \hat{b}_2 \sum X_{1i} X_{2i} \quad \ldots\ldots\ldots (2)$$

ثالثاً – المشتقة الجزئية لـــ $\hat{b}_2$.

$$\frac{\partial \sum \ell_i^2}{\partial \hat{b}_2} = \frac{\partial \sum (Y_i - \hat{b}_0 - \hat{b}_1 X_{1i} - \hat{b}_2 X_{2i})^2}{\partial \hat{b}_2} = 0$$

$$= -2\sum X_{2i}(Y_i - \hat{b}_0 - \hat{b}_1 X_{1i} - \hat{b}_2 X_{2i}) = 0$$

$$\sum X_{2i} Y_i - \sum \hat{b}_0 X_{2i} - \sum \hat{b}_1 X_{1i} X_{2i} - \sum \hat{b}_2 X_{2i} = 0$$

$$\sum X_{2i} Y_i = \hat{b}_0 \sum X_{2i} + \hat{b}_1 \sum X_{1i} X_{2i} + \hat{b}_2 \sum X_{2i}^2 \quad \ldots\ldots (3)$$

إن المعادلات (1)(2)(3) هي المعادلات الثلاثة الطبيعية أو الإعتياديـــة لطريقة المربعات الصغرى .

لدينا القوانين التي يعبر فيها عن المتغيرات بإنحرافاتها عـــن وســـطها الحسابي ، والتي من الممكن أيضاً إستخدامها لإيجاد تقـــديرات الـــمـــــعاملات b_0 و b_1 و b_2 :

$$\hat{b}_1 = \frac{\left[\Sigma\left(X_{1i}-\bar{X}_1\right)\left(Y_i-\bar{Y}\right)\Sigma(X_{2i}-\bar{X}_2)^2\right]-\left[(\Sigma(X_{2i}-\bar{X}_2)(Y_i-\bar{Y}))(\Sigma(X_{1i}-\bar{X}_1)(X_{2i}-\bar{X}_2))\right]}{(\Sigma X_{1i}-\bar{X}_1)^2(\Sigma(X_{2i}-\bar{X}_2)^2)-(\Sigma(X_{1i}-\bar{X}_1)(X_{2i}-\bar{X}_2))}$$

$$\hat{b}_2 = \frac{\left[\Sigma\left(X_{2i}-\bar{X}_2\right)\left(Y_i-\bar{Y}\right)\Sigma(X_{1i}-\bar{X}_1)^2\right]-\left[(\Sigma(X_{1i}-\bar{X}_1)(Y_i-\bar{Y}))(\Sigma(X_{1i}-\bar{X}_1)(X_{2i}-\bar{X}_2))\right]}{(\Sigma X_{1i}-\bar{X}_1)^2(\Sigma(X_{2i}-\bar{X}_2)^2)-(\Sigma(X_{1i}-\bar{X}_1)(X_{2i}-\bar{X}_2))}$$

$$\hat{b}_0 = \bar{Y} - \hat{b}_1\bar{X}_1 - \hat{b}_2\bar{X}_2$$

6-1-2 معامل التحديد أو التقرير

Coefficient of Multiple Determination of Multiple

$R^2y._{X1X2}$ **Coefficient Correlation**

عندما يكون لدينا أكثر من متغير توضيحي في الدالة فإننا نأخذ الإرتباط المتعدد ، إن مربع معامل الإرتباط يدعى معامل التحديد المتعـــدد ، أو مربـــع معامل الإرتباط المتعدد ، إن معامل التحديد أو التقرير المتعدد يرمز له بالحرف الكبير (R^2) مع إشارة الى المتغيرات التي نقوم بدراستها ، مثلاً في نمـــوذج ذو ثلاث متغيرات فإن مربع معامل الإرتباط المتعدد هو ($R^2_{Y.X1X2}$) . وكما فـــي حالة نموذج المتغيرين فإن (R^2) يبين أو يوضح النسبة المئوية للإنحراف الكلي الحاصلة في المتغير المعتمد(Y) الموضحة أو التي توضحها المساحة المستوية

للإنحدار (Regression Plane) أي بوساطة التغيرات الحاصلة في المتغير (X_1) والمتغير (X_2) :

$$R^2{}_{y\cdot x1x2} = \frac{\hat{b}_1 \Sigma (Y_i - \overline{Y})(X_{1i} - \hat{X}_1) + \hat{b}_2 \Sigma (Y_i - \overline{Y})(X_{2i} - \overline{X}_2)}{\Sigma \left(Y_i - \overline{Y}\right)^2}$$

إن قيمة (R^2) تقع بين 0 و 1 فكلما كانت قيمة (R^2) عالية فإن نسبة كبيرة من الإنحرافات في (Y) توضح بوساطة المساحة المستوية للإنحدار ، أي أن التوفيق جيد للمساحة المستوية لمشاهدات العينة الإحصائية ، وكلما كانت قيمة(R^2) أقرب الى الصفر فإن التوفيق أو حسن التطابق يكون رديئاً .

6-1-3 تباين المعلمات $\hat{b}_0$ و $\hat{b}_1$ و $\hat{b}_2$ المقدرة

The Variance of the parameters

يمكن حساب تباين المعلمات المقدرة كما يأتي :

$$Var(\hat{b}_0) = \delta^2 u \left[\frac{1}{n} + \frac{\overline{X}_1{}^2 \Sigma (X_2 - \overline{X}_2)^2 + \overline{X}^2{}_2 \Sigma (X_1 - \overline{X}_1)^2 - 2\overline{X}_1 \overline{X}_2 \Sigma (X_1 - \overline{X}_1)(X_2 - \overline{X}_2)}{\Sigma (X_1 - \overline{X}_1)^2 \Sigma (X_2 - \overline{X}_2)^2 - \left(\Sigma (X_1 - \overline{X}_1)(X_2 - \overline{X}_2)\right)^2} \right]$$

$$Var(\hat{b}_1) = \delta^2 u \left[\frac{\Sigma (X_2 - \overline{X}_2)^2}{\Sigma (X_1 - \overline{X}_1)^2 \Sigma (X_2 - \overline{X}_2)^2 - (\Sigma (X_1 - \overline{X}_1)(X_2 - \overline{X}_2))^2} \right]$$

$$Var(\hat{b}_2) = \delta^2 u \left[\frac{\Sigma (X_1 - \overline{X}_1)^2}{\Sigma (X_1 - \overline{X}_1)^2 \Sigma (X_2 - \overline{X}_2)^2 - (\Sigma (X_1 - \overline{X}_1)(X_2 - \overline{X}_2))^2} \right]$$

عندما يكون :

$$\hat{\delta}_u^2 = \frac{\Sigma \ell i^2}{n - k}$$

حيث أن (k) ترمز الى عدد المعلمات المقدرة ، ففي حالة النمـــوذج ذو الثلاث متغيرات تكون (3=R) ، أما (n) فهي عدد المشاهدات في العينـــة قيـــد الدراسة .

مثال : الجدول (6−1) يحتوي مشاهدات عن الكمية المطلوبة من سلعة معينة (Y) وسعر تلك السلعة (X_1) ودخل المستهلك (X_2) ، قدر دالة الإنحدار لهذه العلاقة :

جدول (1-6)

n	الكمية المطلوبة Yi	X1 السعر	X2 الدخل	$(Y_i-\bar{Y})$	$(X_{1i}-\bar{X}_1)$	$(X_2-\bar{X}_2)$	$(Y_i-\bar{Y})^2$	$(X_{1i}-\bar{X}_1)^2$	$(X_{2i}-\bar{X}_2)^2$	$(Y_i-\bar{Y})(X_{1i}-\bar{X}_1)$	$(Y_i-\bar{Y})(X_{2i}-\bar{X}_2)$	$(X_{1i}-\bar{X}_1)(X_{2i}-\bar{X})$
1	100	5	1000	20	-1	200	400	1	40000	-20	4000	-200
2	75	7	600	-5	1	-200	25	1	40000	-5	1000	-200
3	80	6	1200	0	0	400	0	0	160000	0	0	0
4	70	6	500	-10	0	-300	100	0	90000	0	3000	0
5	50	8	300	-30	2	-500	900	4	250000	-60	15000	-1000
6	65	7	400	-15	1	-400	225	1	160000	-15	6000	-400
7	90	5	1300	10	-1	500	100	1	250000	-10	5000	-500
8	100	4	1100	20	-2	300	400	4	90000	-40	6000	-600
9	110	3	1300	30	-3	500	900	9	250000	-90	15000	-1500
10	60	9	300	-20	3	-500	400	9	250000	-60	10000	-1500
N= 10	ΣY_i = 800	ΣX_1 =60	ΣX_2 =8000	$\Sigma(Y_i-\bar{Y})$ = 0	$\Sigma(X_1-\bar{X}_1)$ = 0	$\Sigma(X_2-\bar{X}_2)$ =0	$\Sigma(Y_i-\bar{Y})^2$ = 3450	$\Sigma(X_1-\bar{X}_1)^2$ = 30	$\Sigma(X_2-\bar{X}_2)^2$ = 1580000	$(X_{1i}-\bar{X}_1)$ = -300	$\Sigma(Y1-\bar{Y})(X_2-\bar{X}_2)$ = 65000	$\Sigma(X_{1i}-\bar{X}_1)(X_{2i}-\bar{X}_2)$ = -5900

$800=\bar{X}_2$ $6=\bar{X}_1$ $80=\bar{Y}_1$

الآن نقوم بما يأتي :

1- نعوض في قانون إيجاد معلمات الإنحدار $\hat{b}_1$ و $\hat{b}_2$:

$$\hat{b}_1 = \frac{(-300)(1580000) - (65000)(-5900)}{(30)(1580000) - (-5900)^2} = \frac{-90500}{12590} = -7.19$$

$$b2 = \frac{(65000)(30) - (-300)(-5900)}{(30)(1580000) - (-5900)} = \frac{180}{12590} = 0.0143$$

لذلك :

$$\hat{b}_0 = \overline{Y} - \hat{b}_1 \overline{X}_1 - \hat{b}_2 \overline{X}_2$$

$$= 80 - (-7.19)(6) - (0.0143)(800) = 111.69$$

2- معامل التحديد أو التقرير المتعدد (R^2) يمكن حسابه على النحو الآتي :

$$R^2 = \frac{\hat{b}_1 \sum (Yi - \overline{Y})(X_{1i} - \overline{X}_1) + (\hat{b}_2 \sum (Y_i - \overline{Y})(X_{2i} - X_2)}{\sum (Y_i - \overline{Y})_2}$$

$$= \frac{(-7.19)(-300) + (0.0143)(65000)}{3450} = 0.894$$

3- لتقدير الخطأ المعياري (The Standard Error) لكل من $\hat{b}_1$ و $\hat{b}_2$:

هنا نحن بحاجة الى تقدير الخطأ المعياري للمتغير العشوائي للمجتمع الإحصائي ($\delta^2{}_u$) . إن الصيغة :

$$\hat{\delta}^2{}_u = \frac{\sum \ell i^2}{n - k}$$

تحتوي على مجموع تربيع البواقي أو الإنحرافات والتي يمكن الحصول عليها من الصيغة الآتية :

$$R^2 = 1 - \frac{\sum \ell^2}{\sum (Y_i - \overline{Y})^2}$$

نقوم بالحل للحصول على ($\sum \ell^2$) نجد :

$$\sum \ell^2 = \sum (Y_i - \overline{Y})^2 (1 - R^2)$$

نعوض كل من ($\sum (Y_i - \overline{Y})^2$) و (R^2) نجد ($\sum \ell^2$) :

$$\sum \ell^2 = (3450)(0.106) = 365.7$$

لذلك :

$$\delta^2_u = \frac{365.7}{10-3} = 52.24$$

4- إن تباين المعلمات $\hat{b}_1$ و $\hat{b}_2$ يمكن الحصول عليه عن طريــق التعــويض بالقانون الآتي :

$$Var(\hat{b}_1) = \delta^2_u \left[\frac{\Sigma(X_{2i}-\bar{X}_2)^2}{\Sigma(X_{1i}-\bar{X}_1)^2 \Sigma(X_{2i}-\bar{X}_2)^2 - (\Sigma(X_{1i}-\bar{X}_1)(X_{2i}-\bar{X}_2))^2} \right]$$

$$Var(\hat{b}_1) = 52.24 \frac{1580000}{12590000} \approx 6.53$$

$$Var(\hat{b}_2) = \hat{\delta}^2_u \left[\frac{\Sigma(X_1-\bar{X}_1)^2}{\Sigma(X_1-\bar{X}_1)^2 \Sigma(X_2-\bar{X}_2)^2 - (\Sigma(X_1-\bar{X}_1)(X_2-\bar{X}_2))^2} \right]$$

$$Var(\hat{b}_2) = 52.24 \frac{30}{12590000} = 0.0001$$

$$Var(\hat{b}_0) = \delta^2_u \left[\frac{1}{n} + \frac{\bar{X}_1^2 \Sigma(X_2-\bar{X}_2)^2 + \bar{X}^2_2 \Sigma(X_1-\bar{X}_1)^2 - 2\bar{X}_1\bar{X}_2 \Sigma(X_1-\bar{X}_1)(X_2-\bar{X}_2)}{\Sigma(X_1-\bar{X}_1)^2 \Sigma(X_2-\bar{X}_2)^2 - (\Sigma(X_1-\bar{X}_1)(X_2-\bar{X}_2))^2} \right]$$

$$Var(\hat{b}_0) = 553.69$$

ولذلك فإن الخطأ المعياري المقدر للمعلمات سيكون بالشكل الآتي :

$$\delta(\hat{b}_1) = 2.55 \qquad \delta(\hat{b}_2) = 0.01 \qquad \delta(\hat{b}_0) = 23.5$$

5- نتائج الإنحدار ربما تعرض الآن على نحو مختصر كما يأتي :

$$\hat{Y} = 111.7 - 7.19X_1 + 0.014X_2$$

$\delta_{(bi)}$	(23.5)	(2.55)	(0.01)	$R^2 = 0.894$
t*	(4.75)	(-2.8)	(1.28)	$t_{0.025} = 2.365$

إن المتغيرين (X_1,X_2) يوضحان 89% من الإنحراف الكلي في (Y) ، أما تقديرات المعلمات ($\hat{b}_0$ $\hat{b}_1$) هي ذات معنوية عالية إحصائياً ولكن تقـــدير ($\hat{b}_2$) هي معلمة ذات معنوية غير عالية في مستوى معنوية 5% .

6-2 نموذج العلاقات غير الخطية في الإنحدار

Nonlinear Relationships in Regression

إن العلاقة الخطية بـــين (Y) والمتغيـــرات التوضـــيحية (X_s) التـــي إفترضناها سابقاً، ربما لاتكون مناسبة لعدد من العلاقـــات الإقتصـــادية ، فـــي الحقيقة إن العلاقة غير الخطية ربما نتوقعها في معظم العلاقات الإقتصـــادية اذا ما عرفنا ان عالم الواقع الحقيقي معقد جداً .

بعض الأشكال المعروفة للعلاقات الإقتصادية غير الخطيـــة يمكـــن أن تعرض بالشكل الآتي مثلاً :

$$Y = b_0 + b_1X_1 + b_2X^2_1 + b_3X^3_1 + \ldots\ldots\ldots\ldots + U$$

أو بوساطة دوال مع مرونات ثابتة مثلاً :

$$Y = b_0X^{b1}_1X^{b2}_2u$$

عندما يكون : b_1 = مرونة ثابتة لــY بالنسبة لــX_1

b_2 = مرونة ثابتة لــY بالنسبة لــX_2

البرهان : بالتعريف فإن المرونة هي :

$$n_{Y\,X1} = \frac{dY}{dX_1} \cdot \frac{X_1}{Y}$$

مشتقة الدالة (Y) بالنسبة للمتغير (X_1) تعطينا :

$$\frac{\partial Y}{\partial X_1} = b_1(b_0 X1^{b1-1} X2^{b2} u) = b_1(b_0 X^{b1}_1 X2^{b2} u)X^{-1}$$

$$= b_1 \frac{Y}{X_1}$$

نعوض عن $\frac{dY}{dX_1}$ بالقيمة الآتية $(b_1 \frac{Y}{X_1})$ في قانون المرونة :

$$n_{Y\,X1} = b_1 \frac{Y}{X_1} \cdot \frac{X_1}{Y} = b_1$$

وهكذا فإن (b_1) هي المرونة الثابتة لـــ(Y) بالنسبة للمتغير (X_1) ، مثلاً في النظرية الإقتصادية الجزئية ، شكل منحنيات معدل الكلفة يأخذ شكل حـــرف (U) يمكن أن يوضع تقريباً بصيغة دالة غير خطية :

$$C = b_0 + b_1X - b_2X^2 + b_3X^3 + U$$

عندما يكون C = الكلفة وX = الإنتاج .

في هذه الحالة معدل الكلفة الكلية هو :

$$\frac{C}{X} = \frac{b_0}{X} + b_1 - b_2X + b_3X^2$$

وهذه الدالة ذات شكل حرف U .

وبالطريقة نفسها فإن دالة الطلب مع مرونة سعرية ثابتة ومرونة دخلية ثابتة يمكن أن يعبر عنها بالمعادلة الآتية :

$$Q_x = b_0 p_x^{b1} Y_u^{b2}$$

عندما يكون : Q_X= الطلب على السلعة X .

P_X= سعر السلعة X .

Y= الدخل تحت التصرف للمستهلكين .

$\frac{dQ}{dP} \cdot \frac{p}{Q} = n_P = b_1$ مرونة الطلب السعرية .

$\frac{dQ}{dY} \cdot \frac{Y}{Q} = n_Y = b_2$ المرونة الدخلية للطلب .

6-3 تحويل الدوال غير الخطية التي تحتوي مرونات ثابتة الى دوال خطية .

اذا كانت العلاقة من نوع العلاقة الثابتة المرونة

$$Y = b_0 X_1^{b1} X_{2}^{b2} u$$

إن الخطأ المعياري (المتغير العشوائي) في هذه الحالة يضــرب بقــيم المتغيرات التوضيحية ، ولذلك لا نستطيع أن نضع الإفتراض الــذي يقــول أن E(u)=0 وذلك لأنه اذا إفترضنا هذا الإفتراض فإن الدالة ستختفي ، بدلاً مــن ذلك نكتب العلاقة التي تحتوي على مرونات ثابتة على نحو ملائم آخر :

$$Y = b_0 X_1^{b1} X_2^{b2} e^{u}$$

عندما يكون e≈2.718 أساس اللوغاريتم الطبيعي ، إن أفضل تحويــل لتقديرات صيغة المرونة الثابتة هو العمل مــع لوغاريتمــات (Logarithms) المتغيرات (الى أساس اللوغاريتم e) بهذه الصيغة :

$$Log_e Y = log_e b_0 + b_1 log_e X_1 + b_2 log_e X_2 + u$$

بعد ذلك نعطي قيمة :

$$log_e X_2 = X^*_2 \text{ , } log_e X_1 = X^*_1 \text{ , } log_e Y = Y^*$$

وبهذا يكون بإمكاننا أن نطبق طريقة المربعات الصــغرى الإعتياديــة (OLS) للدالة المحولة الى العلاقة الخطية :

$$Y^* = b_0^* + b_1 X_1^* + b_2 X_2^* + u$$

مثال :

الجدول (6-2) الآتي يعرض الطلب (Y) على الســلعة (X) وســعرها (X_1) بوحدات لأي عملة ، ونحن نرغب بتقدير دالة الطلب :

جدول (6−2) الطلب على السلعة (X) مع سعرها

المشاهدات	Y الكمية المطلوبة	X_1 السعر	log_eY	$LogX_1$
			Y^*	X^*_1
1	**543**	**61**	**6.2971**	**4.1109**
2	**580**	**54**	**6.3631**	**3.9890**
3	**618**	**50**	**6.4265**	**3.9120**
4	**695**	**43**	**6.5439**	**3.7612**
5	**724**	**38**	**6.5848**	**3.6376**
6	**812**	**36**	**6.6995**	**3.5835**
7	**887**	**28**	**6.7879**	**3.3322**
8	**991**	**23**	**6.8987**	**3.1355**
9	**1186**	**19**	**7.0685**	**2.9445**
10	**1940**	**10**	**7.5705**	**2.3026**

ونحن نرغب في تقدير دالة الطلب الآتية :

$$Y = b_0 X_1^{b1} e^u$$

نقوم بأخذ لوغاريتم المتغيرات (الى أساس اللوغاريتم الطبيعي e) وبعد ذلك نستخدم طريقة المربعات الصغرى الإعتيادية (OLS) في الإنحدار الخطي

$$Log_eY=log_eb_0 + b_1log_eX_1$$

إن البيانات الملاءمة موجودة في الجدول (6−2) ، دوالة الطلب المقدرة هي :

$$log_eY = 9.121 – 0.69\ log_eX \qquad R^2 = 0.992$$

$$\hat{Y} = A . P^{-0.69}$$

(عندما تكون A = مقلوب اللوغاريتم $\hat{b}_0$ = 9.121)

إن مرونة الطلب السعرية هي 0.69− وهذا يعني أن الطلب على السلعة (X) هو طلب غير مرن مع السعر .

الفصل السابع

الإختبارات الإحصائية لمعنوية المعلمات المقدرة بطريقة المربعات الصغرى الإعتيادية

الفصل السابع

الإختبارات الإحصائية لمعنوية المعلمات المقدرة بطريقة المربعات الصغرى الإعتيادية

الإختبارات الإحصائية لمعنوية المعلمات المقدرة بطريقة المربعات الصغرى الإعتيادية

7-1 مقدمـــــــة

في الفصول السابقة تم بناء النموذج والقوانين اللازمة لتقدير معلمــات العلاقات الإقتصادية بإستخدام طريقة المربعات الصغرى الإعتيادية ، والخطوة التي تلي عملية تقدير المعلمات هي الخطوة المتعلقة ببناء معيار للحكــم علــى جودة المعلمات المقدرة ، ثمة ثلاثة معايير لهذا الغرض :

أولاً – المعيار النظري المسبق : وهذا المعيار يوضح من قبل النظرية الإقتصادية ، ويشير الى إشارة وحجوم المعلمات ، فضلاً عن أن هذا المعيــار النظري يحدد في المرحلة الأولى من الدراسة الإقتصادية القياسية وهي مرحلــة تحديد النموذج .

ثانياً – المعيار الإحصائي : وهذا المعيار يوضــح مــن قبــل نظريــة الإحصاء ويتضمن عدداً من الإختبارات الإحصائية لإختبار معنوية المعلمــات المقدرة ، وسيتم التركيز على دراسة هذا المعيار في هذا الفصل .

ثالثاً – المعيار الإقتصادي القياسي : وهذا المعيار يعتمد علــى نظريــة الإقتصاد القياسي .

إن أهم الإختبارات الإحصائية والأكثر إستعمالاً في الإقتصاد القياســي وهما إختبارين أساسيين :

الإختبار الأول : وهو مربع معامل الإرتباط (r^2) الذي يســتخدم للحكــم علــى القابلية التوضيحية للإنحدار الخطــي بــين المتغيــر المعتمــد (Y) والمتغيــر التوضيحي (X) .

الإختبار الثاني : يعتمد على الأخطاء المعيارية (Standard Errors) للمعلمات المقدرة ، ويطبق في حالة درجة الحكم على الإعتماد الإحصائي على المعلمات

المقدرة $\hat{b}_0$ و $\hat{b}_1$ ، إن هذا الإختبار يعطينا مقياس لدرجة الثقة التي يمكــن أن نعزيها أو نعطيها لتقديرات $\hat{b}_0$ و $\hat{b}_1$ ، كما أن هذا الإختبار يمكن الباحــث أن يقرر جودة التقديرات للتعبير عن المعلمــات الحقيقيــة للمجتمــع الإحصــائي (b_1,b_0).

2-7 إختبار جودة توفيق خط الإنحدار بإستخدام مربع معامل الإرتباط (r^2) .

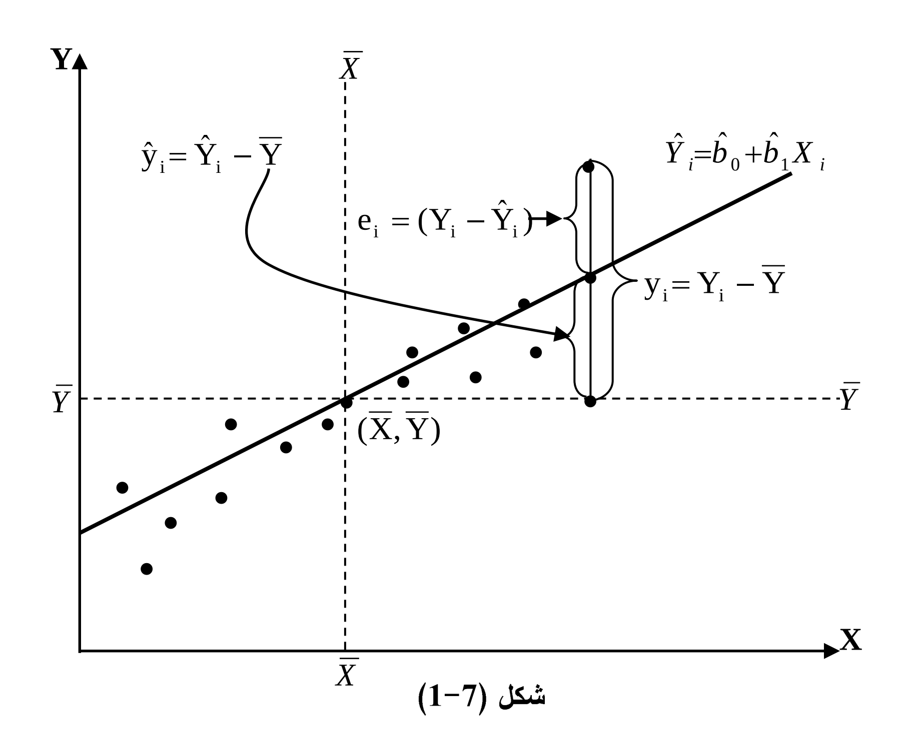
بعد تقدير المعلمات (Parameters) وتحديد خط الإنحــدار للمربعــات الصغرى ، نكون بحاجة الى التعرف أو أن نعرف مدى جــودة توفيــق خــط الإنحدار لمشاهدات العينة الإحصائية من(Y) و(X) وهذا يعني أننا بحاجة الــى أن نقيس إنتشار أو تشتت المشاهدات (dispersion of observations) حول خط الإنحدار ، إن معرفة هذه الحالة يعد ضرورياً وأساسياً ، وذلك لأنه كلمــا كانت المشاهدات قريبة من خط الإنحدار كلما كان ذلك الخط جيد التوفيق وهذا يعني أن خط الإنحدار يوضح على نحو أفضل من الإنحرافــات فــي المتغيــر المعتمد (Y) بسبب التغيرات الحاصلة في المتغير التوضــيحي أو المتغيــرات التوضيحية .

إن مقياس توفيق جودة خط الإنحدار هو(r^2) الذي يبين النسبة المئويــة للإنحراف الكلي (Total Variation) في المتغير المعتمد الذي يمكن توضيحه بوساطة المتغير التوضيحي(X) .

نقوم برسم المشاهدات في شكل الإنتشار وثم نحسب الوســط الحســابي للمتغير التوضيحي والوسط الحسابي للمتغير المعتمد .

$$\bar{X} = \frac{\sum X_i}{n} \qquad \bar{Y} = \frac{\sum Y_i}{n}$$

في الشـــكل (7−1) تبـــدو نقـــاط المشـــاهدات مـــع رســـم الأعمـــدة (prependiculars) التي تمر عبر نقطة الوسط الحســـابي لــــــ(X) والوســـط الحسابي لــــ(Y) .



شكل (7-1)

عند توفيق خط الإنحدار $(\hat{Y}_i = \hat{b}_0 + \hat{b}_1 X_i)$ نحاول أن نحصل على توضيح للإنحرافات الحاصلة في المتغير المعتمد (Y) التي تنتجها أو تسببها التغييرات في المتغير التوضيحي (X) ، وفي الحقيقة إن المشاهدات المنحرفة عن خط الإنحدار المقدر تبين أن خط الإنحدار يوضح جزءاً من الإنحراف الكلي في المتغير المعتمد فقط ، فهناك جزء من انحراف يعبر عنه بالشكل الآتي :

$$e_i Y_i - \hat{Y}_i$$

يبقى بدون توضيح .

1- ربما يكون بإمكاننا أن نحسب الإنحراف الكلي للمتغير المعتمد (Y) عن طريق مقارنة كل قيمة لـ(Y) مع قيمة الوسط الحسابي لـ(Y) التي هي ($\overline{Y}$) ونضيف كل النتائج للإنحرافات ، نرمز للإنحرافات في قيمة (Y_i) عن الوسط الحسابي بالحرف الصغير (y) لذلك يكون لدينا :

$$\sum_{i=1}^{n} y_i^2 = \sum_{i=1}^{n} (Y_i - \overline{Y})^2$$ الإنحراف الكلي في قيم (Y)

يجب أن نلاحظ أنه من أجل الحصول على الإنحراف الكلي في قيم (Y) يجب علينا أن نربع الإنحرافات عن الوسط الحسابي ، وذلك لأن مجموع الإنحرافات البسيطة عن الوسط الحسابي لأي متغير تساوي صفراً ، وكما يأتي:

$$\sum_{i=1}^{n} (Y_i - \overline{Y}) = \sum_{i=1}^{n} y_i = 0$$

2- بالطريقة نفسها تعرف الإنحرافات للقيم المقدرة من خط الإنحدار ($\hat{Y}$) عن الوسط الحسابي ($(\hat{y}_i = \hat{Y}_i - \overline{Y})$) ، إن هذا الجزء من الإنحراف الكلي لقيم (Y_i) التي يوضحها خط . وهكذا فإن مجموع تربيع هذه الإنحرافات هو مجموع الإنحرافات في المتغير المعتمد التي يوضحها خط الإنحدار .

$$\sum_{i=1}^{n} \hat{y}_i^2 = \sum_{i=1}^{n} (\hat{Y}_i - \overline{Y})^2$$

3- عرفنا الإنحراف في القيم الفعلية عن القيم المقدرة بالقيمة ($e_i = Y_i - \hat{Y}_i$) وهو الجزء من الإنحرافات في المتغير المعتمد التي لا يوضحها خط الإنحدار والتي تعزى الى الإضطراب الذي يسببه المتغير العشوائي (U) ، وهكذا فإن مجموع تربيع هذه الإنحرافات يعطينا مجموع الإنحراف في المتغير المعتمد (Y) الذي لا يوضحه خط الإنحدار ، أي مجموع تربيع الإنحراف غير الموضح

$$\sum_{i=1}^{n} \ell_i^2 = \sum_{i=1}^{n} (Y_i - \hat{Y})^2$$ = الإنحراف غير الموضح

وعلى نحو مختصر :

إنحراف المشاهدات (Y_i) عن خط الإنحدار $e_i = Y_i - \hat{Y}_i$

إنحراف المشاهدات (Y_i) عن وسطها الحسابي $y_i = \hat{Y}_i - \overline{Y}$

إنحراف القيم المقدرة (Y_i) عن الوسط الحسابي $\hat{y}_i = \hat{Y}_i - \bar{Y})$

$$\sum y_i^2 = \sum \hat{y}_i^2 + \sum \ell_i^2$$

$$\begin{pmatrix}\text{الإنحراف}\\ \text{الكلي}\end{pmatrix} = \begin{pmatrix}\text{الإنحراف الذي يوضحه}\\ \text{خط الإنحدار}\end{pmatrix} + \begin{pmatrix}\text{الإنحراف غير الموضح}\\ \text{بوساطة خط الإنحدار}\end{pmatrix}$$

$$r^2 = \frac{\sum \hat{y}_i^2}{\sum y_i^2} = \frac{\text{الإنحراف الموضح}}{\text{الإنحراف الكلي}}$$

(r^2) يقرر النسبة من الإنحرافات في المتغير المعتمد (Y) التي يوضحا الإنحراف في قيم المتغير التوضيحي (X) ، لهذا السبب فإن (r^2) في بعض الأحيان يدعى معامل التحديد أو التقرير (Coefficient of Determination).

مثلاً اذا كان ($0.90 = r^2_{YX}$) فإن هذا يعني إن خط الإنحدار يعطينا توفيق جيد للبيانات المشاهدة وذلك لأن هذا الخط يوضح (90%) من الإنحراف الكلي في قيم المتغير المعتمد (Y) عن وسطها الحسابي ، أما نسبة (الـ10%) المتبقية من الإنحراف الكلي في (Y) لا يوضحها خط الإنحدار ، ولكنها تعزى الى العوامل الداخلة في المتغير العشوائي الذي يسبب الإضطراب (Disturbance) .

إن قيم معامل التحديد (r^2) يمكن أن تأخذ قيماً تقع بين الصفر والواحد ، بعبارة أخرى إن (r^2) يأخذ قيماً على النحو الآتي :

$$0 \leq r^2 \leq 1$$

كما أن هناك علاقة بين مربع معامل الإرتباط الذي هو (r^2) وميل الدالة أو ميل خط الإنحدار المقدر يمكن إيجاده من خلال القانون الآتي :

$$r^2 = \hat{b}_1 \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sum (Y_i - \bar{Y})^2}$$

مثلاً : في مثالنا عن دالة العرض في فصل سابق قدرنا دالة العـــرض كما يأتي :

$$\hat{Y} = 33.75+3.25X_i$$

ومن البيانات وجدنا أن :

$$156 = \sum(X_i - \overline{X})(Y_i - \overline{Y})$$
$$894 = \sum(Y_i - \overline{Y})$$

وبإستخدام القانون أعلاه لإيجاد (r^2) عن طريق ($\hat{b}_1$) التي هي (3.25) المعطاة في الدالة المقدرة :

$$r^2 = 3.25 \times \frac{156}{894}$$
$$= 0.570$$

3-7 إختبار الخطأ المعياري للمعلمات المقدرة ($\hat{b}_0$ و $\hat{b}_1$) بطريقة المربعـات الصغرى الإعتيادية .

إن المعلمات المقدرة بطريقة المربعات الصـــغرى ($\hat{b}_0$ و $\hat{b}_1$) يمكـــن الحصول عليها من عينة إحصائية (Sample) عن المشاهدات لمتغير المعتمـــد (Y) والمتغير التوضيحي (X) ، ولما كانت الأخطـــاء فـــي تقـــديرات العينـــة الإحصائية مسألة حتمية في كل المعلمات المقدرة أصبح من الضروري تطبيـــق إختبارات المعنوية الإحصائية من أجل قياس حجم الخطأ وتحديد درجة الثقة في مشروعية المعلمات المقدرة ، هناك إختبارات عدة لهذا الغرض ، فـــي الفقـــرة الحالية من هذا الفصل سنقوم بدراسة واحداً من الإختبارات المهمة وهو "إختبار الخطأ المعياري" الذي يعد شائعاً في البحث الإقتصادي القياسي التطبيقـــي ، إن هذا الإختبار يساعدنا أن نقرر كون المعلمات المقدرة ($\hat{b}_0$ و $\hat{b}_1$) هي معلمات

مختلفة عن الصفر على نحو كبير ، وهذا يعني كون العينة الإحصائية التي تم من خلالها الحصول على تلك التقديرات للمعلمات قد جاءت من مجتمع إحصائي معلماته الحقيقية لا تساوي صفراً ، وعلى نحو منهجي نختبر الفرضيتين الآتيتين:

أولاً : فرضية العدم (Null Hypothesis)

$$H_o : b_i = 0$$

ثانياً : الفرضية البديلة (Alternative Hypothesis)

$$H_1 : b_i \neq 0$$

يمكن أن يلخص إختبار الخطأ المعياري أو الإنحراف المعياري بالشكل الآتي :

$$\delta(\hat{b}_1) = \sqrt{\frac{\sum \ell_i^2}{(n-2)\left[\sum \left(X_i - \overline{X}\right)^2\right]}}$$

$$\delta(\hat{b}_0) = \sqrt{\frac{\left(\sum \ell_i^2\right)\sum X_i^2}{(n-2)n\sum (X_i - \overline{X})^2}}$$

بعد ذلك نقارن الإنحرافات المعيارية مع القيم الرقمية للمعلمات المقدرة ($\hat{b}_0$ و $\hat{b}_1$) ، فإذا كان الإنحراف المعياري هو أصغر من نصف القيمة الرقمية للمعلمة المقدرة ($\delta(\hat{b}_1) < \frac{\hat{b}_1}{2}$) فإننا نخلص الى القول إن تلك التقديرات هي إحصائياً ذات معنوية إحصائية ، وهذا يعني أننا نرفض فرضية العدم التي تنص على أن المعلمات الحقيقية للمجتمع الإحصائي ($b_i=0$) ، وهذا يساوي قبول الفرضية البديلة التي تنص على أن المعلمات الحقيقية للمجتمع الإحصائي (b_i) مختلفة عن الصفر ، من جانب آخر اذا كان الخطأ المعياري للمعلمات المقدرة

هو أكبر من نصف القيمة الرقمية ($\delta(\hat{b}_i)\langle\frac{\hat{b}_1}{2}$) والنتيجة هي أن المعلمات المقدرة ليست ذات معنوية إحصائية ، وهذا يعني أننا نقبل فرضية العدم التي تنص على ان المعلمة الحقيقية للمجتمع الإحصائي ($b_i=0$) .

ومن أجل تسهيل عملية المقارنة بين الأخطاء المعيارية للمعلمات المقدرة مع القيم الرقمية للمعلمات المقدرة ، فإنه من الملائم أن نضع الأخطاء المعيارية بين اقواس تحت المعلمات المقدرة التي تشير اليها تلك الأخطاء المعيارية .

مثال : الأخطاء المعيارية لمعلمات دالة العرض التي تم تقديرها في فصول سابقة هي كما يأتي :

$$\delta(\hat{b}_1)=\sqrt{\frac{\sum \ell_i^2}{(n-2)\left[\sum\left(X_i-\overline{X}\right)^2\right]}}=\sqrt{\frac{383.98}{(10)(48)}}\approx 0.9$$

$$\delta(\hat{b}_0)=\sqrt{\frac{\left(\sum \ell_i^2\right)\sum X_i^2}{(n-2)n\sum(X_i-\overline{X})^2}}=\sqrt{\frac{(383.98)(1020)}{(120)(48)}}\approx 8.3$$

والآن يمكن أن تعرض النتائج للإنحدار بصيغة دالة :

$$\hat{Y}_i=33.75+3.25X_i$$

$$(8.3)\qquad(0.9)$$

على هذا النحو يمكن إجراء إختبار سريع لمعنوية المعلمات المقدرة عن طريق الفحص :

$$\delta(\hat{b}_1)\langle\frac{\hat{b}_1}{2}\qquad\qquad \delta(\hat{b}_0)\langle\frac{\hat{b}_0}{2}$$

وهكذا فإن ($\hat{b}_0$ و $\hat{b}_1$) هما ذات معنوية إحصائية عاليـــة ، أي أنهمـــا مختلفان عن قيمة الصفر ، وهذا يعني أننا نرفض فرضية العدم ونقبل الفرضية البديلة .

7-4 إختبار (Z) للمعلمات المقدرة بطريقة المربعات الصغرى

إن هذا الإختبار يعتمد على التوزيع الطبيعي ، يطبق هذا الإختبار فـــي الحالتين الآتيتين :

أ- اذا كان تباين المجتمع الإحصائي معروفاً .

ب- اذا كان تباين المجتمع الإحصائي غير معروف .

فضلاً عن ان العينة الإحصائية التي يتم بوساطتها الوصول الى تقديرات المعلمات للمجتمع الإحصائي هو حجم كبير ما فيه الكفايه ($n>30$) أكثـــر مـــن ثلاثين مشاهدة ، فإذا لم تكن هذه الشروط موجودة فإننا نلجأ الى تطبيق إختبـــار (t) الذي سوف نوضحه في الفقرة القادمة .

في تطبيقات الإقتصاد القياسي فإن تباين المجتمع الإحصائي لـــ(Y) هو تباين (u) الذي هو (δ_u^2) وهو مجهول ، ولكن اذا كانت لدينا عينة إحصـــائية كبيرة ($n>30$) فإننا نستطيع إستخدام التوزيع الطبيعي ونطبق إختبار (Z) على نحو تقريبي ، وذلك لأن التباين المقدر للعينة الإحصائية (δ^2) هو عبارة عـــن تقدير تقريبي لتباين المجتمع الإحصائي المجهول (δ_u^2) بالنسبة لعينة كبيـــرة الحجم .

إن إختبار (Z) ربما يوضح بالخطوات الآتية :

نريد أن نختبر أولاً فرضية العدم (Null Hypothesis)

$$\mathbf{H_o : b_i = 0}$$

مقابل أو ضد الفرضية البديلة (Alternative Hypothesis)

$H_1 : b_i \neq 0$

قلنا سابقاً أنه في ظل ظروف فرضيات معينة تتعلق بقيم المتغيــر (U) فإن (U) سيتبع نظام ([$U \sim N(0, \delta_u^2)$]) ، أي أن المتغيــر (U) لــه وســط حسابي يساوي صفراً وإن له تباين هو (δ_u^2) فإن تقديرات المربعات الصغرى ($\hat{b}_0$ و $\hat{b}_1$) لهما التوزيعات الإعتيادية الآتية :

$$\hat{b}_0 \sim N\left(b_0, \delta_{(\hat{b}_0)} = \sqrt{\delta_u^2 \frac{\sum X^2}{n \sum (X_i - \overline{X})^2}} \right)$$

$$\hat{b}_1 \sim N\left(b_1, \delta_{(\hat{b}_1)} = \sqrt{\delta_u^2 \frac{1}{n \sum (X_i - \overline{X})^2}} \right)$$

إن التوزيعات الطبيعية أعلاه يمكن أن يقاس عليها أو إعتمادها بوصفها معياراً ، أي يمكن تحويلها الى وحدات متغير معياري (Z) الذي يتميز بكــون وسطه الحسابي يساوي صفراً وتباين (($Z \sim N(0,1)$) من خلال قانون التحويل :

$$Z_i = \frac{X_i - \mu}{\delta} \sim N(0,1)$$

عندما يكون :

(X_i) قيمة المتغير الذي نريد أن نجعله طبيعياً (أي تحويل قيمة المتغير (X) الى وحدات (Z) المعيارية) .

(μ) الوسط الحسابي لتوزيع المتغير .

(δ) الإنحراف المعياري للمتغير .

في حالة توزيع تقديرات المربعات الصغرى ($\hat{b}_0$ و $\hat{b}_1$) فــإن قــانون التحويل أعلاه يأخذ الشكل الآتي :

بالنسبة الى $\hat{b}_0$

$$Z = \frac{\hat{b}_0 - b_0}{\delta_{(\hat{b}_0)}} = \frac{\hat{b}_0 - b_0}{\sqrt{\frac{\delta_u{}^2 \sum X_i{}^2}{n\sum(X_i - \overline{X})^2}}} \sim N(0,1)$$

بالنسبة الى $\hat{b}_1$

$$Z = \frac{\hat{b}_1 - b_1}{\delta_{(\hat{b}_1)}} = \frac{\hat{b}_1 - b_1}{\sqrt{\frac{\delta_u{}^2}{\sum(X_i - \overline{X})^2}}} \sim N(0,1)$$

بإستخدام قوانين التحويل أعلاه نستطيع إجراء إختبار لأية فرضية تتعلق بالقيمة الحقيقية للمعلمة ($\hat{b}_i$) للمجتمع الإحصائي ، نفترض أننا نريد أن نختبر فرضية العدم التي تنص على أن القيمة الحقيقية للمعلمة ($\hat{b}_1$) هي مساوية الى قيمة معينة مثلاً(b^*_1) ، وهنا فإن غرضنا الأساسي هو إختبار فرضية العدم :

$$H_o : b_1 = b^*_1$$

مقابل أو ضد الفرضية البديلة

$$H_1 : b_1 \neq b^*_1$$

نعوض عن (b_1) بقيمة (b^*_1) في القانون أعلاه ، ولدينا تقدير قيمة ($\hat{b}_1$) والإنحراف المعياري أو الخطأ المعياري للمعلمة (b_1) هو ($\delta_{(\hat{b}_1)}$)

وبذلك نستطيع حساب قيمة (Z^*) :

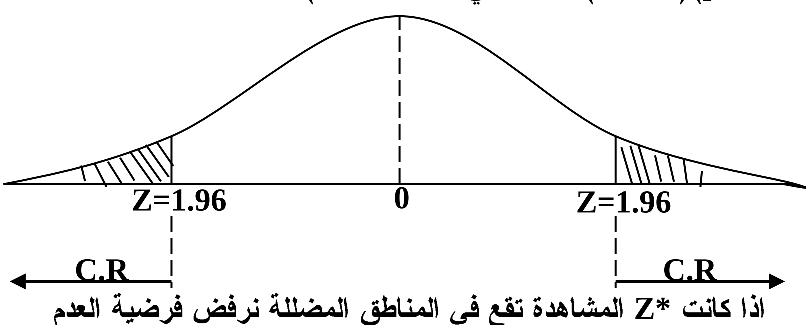
$$Z^* = \frac{\hat{b}_1 - b_1{}^*}{\delta(\hat{b}_1)}$$

بعد أن حصلنا على هذه القيم نستطيع حساب إحتمال الحصـــول علـــى تقدير لـــ(b_1)اذا كانت فرضيتنا الأساسية($b_1 = b^*_1$) هي حقيقية نتبع الخطـــوات الآتية :

نختار مستوى من مستويات المعنوية من أجل أن نقرر قبول أو رفض الفرضية (إن مستوى المعنوية هو إحتمال إتخاذ القرار الخطأ يعني إحتمال رفض الفرضية عندما تكون فعلياً هي الصحيحة أو الحقيقية) .

عادة في البحث الإقتصادي القياسي نختار مستوى (5%) أو (1%) من المعنوية ، وهذا يعني أننا في حالة إتخاذنا قراراً نسمح لإحتمال (5%) أن يكون قرارنا خطأ أي قرار رفضنا للفرضية عندما تكون فعلياً هي الصحيحة ، وعادة في العمل الإقتصادي القياسي التطبيقي يستخدم إختبار ذو نهايتين أو جانبين (Tow-Tail Test) ، إن إختيار إختبار الجانبين يتضمن عدم وجود معرفة مسبقة تتعلق بإشارة المعلمة التي يراد إختبار معنويتها الإحصائية ، ولكن على كل حال فإن إختيار الجانب الواحد أو النهاية الواحدة (One-Tail Test) سيكون أكثر ملائمة في أغلبية التطبيقات الإقتصادية القياسية وذلك لأن النظرية الإقتصادية تقوم بإعطائنا معلومات أو توقعات مسبقة تتعلق بإشارة المعلمات للعلاقات الإقتصادية .

يجب أن نختار المنطقة الحرجة (Critical Region) ، بحيث أن إختيار كلا الجانبين للتوزيع المعياري الطبيعي وبخاصة ذلك الجزء لكل جانب من الإختبار الذي يقابل نصف إحتمال المعنوية ، مثلاً اذا إخترنا مستوى معنوية (5%) ، فإن كل جانب سيحتوي المنطقة أو الجزء من الإحتمال (probability)(0.025) ، كما في الشكل (7−2) :



شكل (7−2) إختبار ذو نهايتين أو جانبين مع 5% مستوى معنوية

من جدول التوزيع المعياري الطبيعي نستطيع أن نجد القيم الحرجة لـ(Z) التي تقابل الإحتمال (0.025) في كل نهاية من نهايات المنحنى .

($1.96 = Z_2$, $-1.96 = Z_1$) والخطوة الأخيرة هي مقارنة قيمة (Z^*) المشاهدة مع القيم الحرجة لـ(Z) .

فإذا كانت (Z^*) المشاهدة تقع ضمن المنطقة الحرجة أي أن ($Z^*<1.96$ أو $1.96<Z^*$) فإننا نرفض فرضية العدم التي تقول أن القيمة الحقيقية لـ(b) هي (b^*) وذلك لإن إحتمال (Z^*) صغير جداً .

بعبارة أخرى ليس من المحتمل أن مثل (Z^*) يمكن ان تشاهد اذا كانت الفرضية الأساسية (H_0) هي الحقيقية أو الصحيحة ، فإذا كان الوضع بالعكس فإن قيمة (Z^*) للعينة يقع خارج المنطقة الحرجة المختارة ، وهذا يعني أن ($-1.96<Z^*<1.96$) فإننا نقبل فرضية العدم($H_o : b_1 = b^*_1$) وذلك لإن إحتمال أن يكون (Z^*) اذا كانت فرضية العدم هي الحقيقية كبيراً .

مثال (1) : نفترض أن $\hat{b}_1 = 29.48$ و $\delta_{(\hat{b}_1)} = 36.0$

ونريد أن نختبر فرضية العدم (H_0)

$$H_o : b^*_1 = 25.0$$

من قانون التحويل لـ(Z) نحصل على :

$$Z^* = \frac{\hat{b}_1 - b_1^*}{\delta_{(\hat{b}_1)}} = \frac{29.48 - 25.0}{36.0} = 0.12$$

لإن (Z^*) لم تقع في المنطقة الحرجة ($Z^*<1.96$) فإننا نقبل فرضية العدم بإن ($b^*=25.0$) وذلك لأن إحتمال المشاهدة لمثل هذه القيمة (Z^*) (أكبر من 5%) .

ولكن في الإقتصاد القياسي التطبيقي أصبح معتاداً لإختبار الفرضية التي تقول أن المعلمة الحقيقية للمجتمع الإحصائي تساوي صفراً ، وهذا هو الشكل النموذجي لفرضية العدم في الإقتصاد القياسي وهي :

$$H_o : b_i = 0$$

ويتم إختبار هذه الفرضية ضد أو مقابل الفرضية البديلة :

$$H_1 : b_i \neq 0$$

إن معنى ومتضمنات هذه الفرضية قد تم شرحه في الفقرة السابقة ولكن يمكن أن نلخص المناقشة كما يأتي :

اذا رفضنا فرضية العدم ، فإن المعلمة التجريبية ($\hat{b}_1$) هي ذات معنوية إحصائية عالية ، وهذا يعني أنها مختلفة عن الصفر على نحو كبير أو مهم .

أما اذا قبلنا فرضية العدم فإن هذا يعني أن المعلمة ($\hat{b}_1$) هي ليست ذات معنوية إحصائية وأنه ربما لا تكون هناك علاقة خطية بين (X) و (Y) في المجتمع الإحصائي .

من أجل إجراء إختبار فرضية العدم أعلاه نجعل (b=0) في قانون التحويل الى قيم التوزيع الطبيعي في (Z) :

$$Z^* = \frac{\hat{b}_1 - b_1^*}{\delta_{(\hat{b}_1)}} = \frac{\hat{b}_1 - 0}{\delta_{(\hat{b}_1)}} = \frac{\hat{b}_1}{\delta_{(\hat{b}_1)}}$$

وهكذا في حالة إختبار فرضية العدم($H_o : b_i = 0$) فإن أسلوب إختبار (Z) يختصر الى الخطوة البسيطة بتقسيم القيمة المقدرة للمعلمة (b_1) على الإنحراف المعياري أو الخطأ المعياري لتلك القيمة وبعدها مقارنة (Z^*) الناتجة من التقسيم مع القيمة النظرية (الجدولية) لـ(Z) والتي تعطينا المنطقة الحرجة للإختبار ، القيمة النظرية لـ(Z) يمكن الحصول عليها من جدول التوزيع الطبيعي .

مثال (2) : نفترض أننا قدرنا دالة العرض من عينة إحصائية متكونة من (700) مشاهدة أي أن n=700 ، والدالة المقدرة كانت كما يأتي :

$$\hat{Y}_i = 100 + 4.00X_i$$
$$(20) \quad (1.5)$$

والآن سنجري إختبار (Z) للميل المقدر ($\hat{b}_1$=4) ، الذي له خطأ معيارياً مساوياً الى (1.5) ، طالما أن العينة الإحصائية كبيرة الحجم ، فإن الإنحراف المعياري المقدر للمعلمات هو تقريب جيد للإنحراف الحقيقي لهذه المعلمات في المجتمع الإحصائي ، لذلك نستطيع تطبيق إختبار (Z) لإيجاد المعنوية الإحصائية لتقديرات ($\hat{b}_0$ و $\hat{b}_1$) :

فرضية العدم : $H_0 : b_1 = 0$

الفرضية البديلة : $H_1 : b_1 \neq 0$

الآن نحسب قيمة (Z^*) نجد أن :

$$Z^* = \frac{\hat{b}_1}{\delta(\hat{b}_1)} = \frac{4}{1.5} = 2.66$$

إن القيمة النظرية (الجدولية) لـ(Z) (عند مستوى 5% من المعنوية) هي 1.96 ، وهذا يعني ($Z<Z^*$) ولذلك نرفض فرضية العدم التي تقول أن (b_1=0) ، ونقبل الفرضية البديلة ، وهذا يعني أن الإنحدار المقدر ذا معنوية إحصائية عالية .

7-5 إختبار (t) للمعلمات المقدرة بطريقة المربعات الصغرى

إن هذا الإختبار مشابه لإختبار (Z) إلا إنه ملائم عندما يكون حجم العينة الإحصائية صغيراً (أقل من ثلاثين مشاهدة) ، ولإجراء إختبار (t) ذو النهايتين أو الجانبين نتبع الخطوات الآتية :

أ- نحدد فرضية العدم والفرضية البديلة .

ب- نختار مستوى المعنوية المرغوب (الدلالة) مثلاً 5% أو 1% .

ج- نحدد عدد درجات الحرية (Degree of Freedom) .

عندما تكون هذه المعلومات موجودة فمن الممكـــن أن نحـــدد المنطقـــة الحرجة (Critical Region) والتي تفصل بين مناطق الرفض ومناطق القبول، أما القانون الذي بموجبه يمكن إيجاد قيمة (t) هو :

$$t^* = \frac{\hat{b}_i - b_1^*}{\sqrt{\hat{\delta}^2_{(\hat{b}_1)}}}$$

عندما يكون :

$(\hat{b}_i)$ = تقدير (b_i) من المربعات الصغرى .

(b^*_i) = قيمة إفتراضية لـــ(b_i) $(\mathbf{H_o : b_i = b^*_1})$.

$(\hat{\delta}^2_{(\hat{b}_1)})$ = التباين المقدر لـــ(b_i) من الإنحدار .

(n) = حجم العينة الإحصائية .

(k) = العدد الكلي للمعلمات المقدرة .

عادةً في الإقتصاد القياسي يتم وضع أو بناء فرضية العدم أولاً :

$$\mathbf{H_o : b_i = 0}$$

ويتم إختيار هذه الفرضية ضد الفرضية البديلة :

$$\mathbf{H_1 : b_i \neq 0}$$

في هذه الحالة فإن (t) يقلص الى ما يأتي :

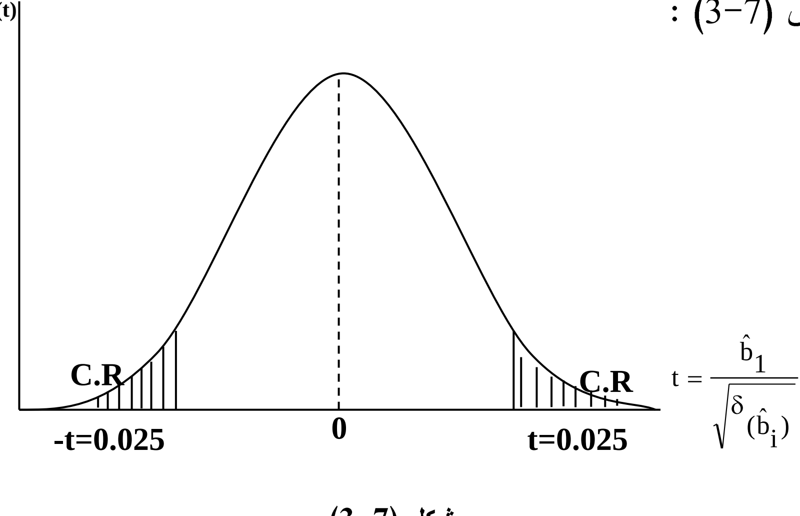
$$t^* = \frac{\hat{b}_i}{\hat{\delta}_{(\hat{b}_i)}}$$

إن قيمة (t^*) للعينة تقدر بتقسيم القيمة المقدرة للمعلمة على الخطأ المعياري (Standard Error) ، وهذه القيمة تقارن مع القيمة النظرية أو الجدولية لــ(t) والتي عن طريقها تحدد المنطقة الحرجة في إختبار النهايتين أو الجانبين مع درجات حرية مساوية لــ(n-k) .

1- اذا وقعت (t^*) في المنطقة الحرجة فإننا نرفض فرضية العدم أي نقبل بإن تقدير ($\hat{b}_i$) هو ذا معنومية إحصائية عالية .

2- اذا وقعت (t^*) في منطقة القبول (acceptance region) عندئذ نقبل فرضية العدم بإن تقدير ($\hat{b}_i$) هو ليس ذا معنوية إحصائية عالية بمستوى معنوية 5% .

إن إختبار النهايتين لإختبار فرضية العدم (**$H_o : b_i = 0$**) موضحاً في الشكل (3-7) :



شكل (3-7)

فإذا كانت قيمة (t) المشاهدة (t^*) تقع في المنطقة الحرجة (المنطقة المضللة) فيجب أن نرفض فرضية العدم عند مستوى معنوية (0.05) .

الفصل السابع

الإختبارات الإحصائية لمعنوية المعلمات المقدرة بطريقة المربعات الصغرى الإعتيادية

مثال (1) : إفترض اننا حصلنا من عينةذات حجم (n=20) على تقــدير لدالــة الإستهلاك :

$$\hat{C} = 100 + 0.70Y$$
$$\mathbf{(75.5)} \quad \mathbf{(0.21)}$$

إن الأرقــام بيــت الأقــواس هــي الأخطــاء المعياريــة لمعلمــات $(\hat{b}_0 = 100)(\hat{b}_1 = 0.70)$ ، لإن حجم العينة أقل من (30) مشاهدة فليس بالإمكان إســتعمال إختبــار (Z) ، ولكــن نحــن نعلــم الإفتراضــات الإحتماليــة (Stochastic Assumptions) حول قيم المتغير العشوائي (U) فإن التقديرات موزعة توزيعاً طبيعياً ، ولذلك نستطيع إختبار (t) نجد أن :

$$t^* = \frac{\hat{b}_i}{\hat{\delta}_{(\hat{b}_i)}} = \frac{0.70}{0.21} \approx 3.3$$

نريد أن نختبر الفرضية (العدم) كما يأتي :

$$\mathbf{H_0 : b_1 = 0}$$

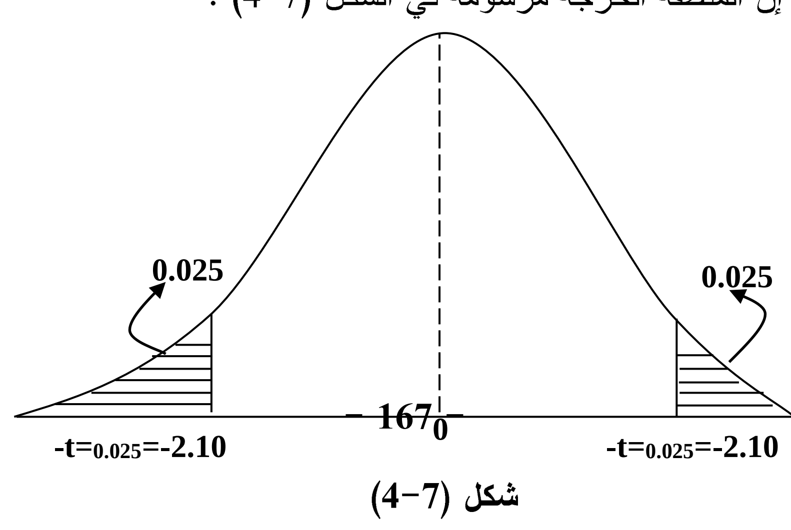
ضد الفرضية البديلة :

$$\mathbf{H_1 = b_1 \neq 0}$$

القيم الحرجة لــ(t) بالنسبة الى (18=n-k) مقابل هذا نجد قيمة (t) :

$$\mathbf{t_1 = -t_{0.025} = -2.10}$$
$$\mathbf{t_2 = +t_{0.025} = +2.10}$$

إن المنطقة الحرجة مرسومة في الشكل (7−4) :



شكل (4−7)

لإن ($t^*=3.3$) أكبر من القيمة الجدولية لـــ(t) (2.10) فـــي مســـتوى (0.025) ، فإننا نرفض الفرضية العدمية ونقول أن (b_1) هـــي مختلفـــة عـــن الصفر .

مثال (2) : اذا كانت لدينا المعلومات الآتية عن دالة العرض :

$$\hat{Y} = 33.75 + 3.25X$$

$$\delta_{(\hat{b}_0)} = \sqrt{\delta^2 u \frac{\sum X_i^2}{n\sum(X_i - \overline{X})^2}} = 8.28$$

$$\delta_{(\hat{b}_1)} = \sqrt{\delta_u^2 \frac{1}{\sum(X_i - \overline{X})^2}} = 0.89$$

الآن نجد قيمة (t^*) لكل من($\hat{b}_0$ و $\hat{b}_1$) :

$$t^*(\hat{b}_0) = \frac{\hat{b}_0}{\hat{\delta}_{(\hat{b}_0)}} = 4.07$$

$$t^*(\hat{b}_1) = \frac{\hat{b}_i}{\hat{\delta}_{(\hat{b}_i)}} = 3.62$$

ومن الواضح ان كلا القيمتين لـــ(t^*) هي عالية وهذا يعني أن التقديرات هي إحصائياً ذات معنوية عالية .

7-6 حدود الثقة للمعلمات b_0 و b_1

إن رفض فرضية العدم (Null Hypothesis) لا يعنـــي أن تقـــديراتنا لـــ($\hat{b}_1$) هي تقديرات صحيحة للمعلمات الإحصائية للمجتمع الإحصائي ، ولكن ذلك يعني أن تقديراتنا للمعلمات قد جاءت من عينة إحصائية أخذت من مجتمع إحصائي معلماته تختلف عن الصفر ، من أجل أن نرى الى أي مـــدى تقتـــرب تقديراتنا لـــ(b_i) من المعلمات الحقيقية (True parameters) يجب أن نضـــع حدود الثقة (Confidence intervals) للمعلمات الحقيقية ، بعبارة أخرى نقوم بوضع قيم محددة تكون ضمنها تقديرات المعلمات ، لذلك نقول أنـــه بمســـتوى معطى من الإحتمالية ، فإن معلمة المجتمع الإحصائي سيكون ضمن حدود الثقة المذكورة أو المحددة .

نقوم بإختيار نسبة الإحتمال (a probability) مسبقاً ونشـــير اليهـــا ، وندعوها مستوى الثقة (Confidence Level) . لقد أصبح من المعتاد في علم الإقتصاد القياسي إختيار 95% مستوى ثقة ، وهذا يعني أنه في حالة إعادة أخذ العينات فإن حدود الثقة (Confidence limits) والمحتسبة من العينة ســـوف تتضمن أو تحتوي على المعلمة الحقيقية للمجتمع الإحصـــائي فـــي 95% مـــن الحالات وهناك 5% من الحالات تكون فيها المعلمة الحقيقية للمجتمع الإحصائي خارج حدود الثقة .

7-6-1 حدود الثقة من إختبار (Z)

لقد ذكرنا سابقاً إنه بالإمكان إستخدام توزيع (Z) فـــي حالـــة معرفـــة الإنحراف المعياري للمجتمع الإحصائي ($\delta_{(\hat{b})}$) ، أو في حالـــة تـــوفر عينـــة إحصائية كبيرة (n أكبر من 30) ، وذلك لإن العينات الكبيـــرة لهـــا إنحـــراف معياري (δ) يعد تقديراً جيداً للإنحراف المعياري للمجتمع الإحصائي المجهول

إن (Z) للمعلمات ($\hat{b}_i$) هو :

$$Z = \frac{\hat{b}_i - b_i}{\delta_{(\hat{b}_i)}}$$

إن واجبنا الأول هنا هو إختيار مستوى الثقة ولنقل 95% ، بعدها ننظر الى الجدول النظري (Standard normal table) ونجد أن إحتمال قيمـــة (Z) تقع بين (‎−1.96)و(‎+1.96) في مستوى الثقة 95% وهذا يمكن التعبير عنه كما يأتي :

$$P\{-1.96 \langle Z \langle +1.96 = 0.95\}$$

نعوض عن قيمة (Z) :

$$P\left\{-1.96 \langle \frac{\hat{b}_1 - b_i}{\delta_{(\hat{b}_i)}} \langle 1.96\right\} = 0.95$$

$$P\left\{\hat{b}_i - 1.96(\delta_{(\hat{b}_i)}) \langle b_i \langle \hat{b}_i + 1.96(\delta_{(\hat{b}_i)})\right\} = 0.95$$

وهكذا فإن حدود الثقة بمستوى 95% للمعلمة (b_i) هي :

$$\left\{\hat{b}_i - 1.96(\delta_{(\hat{b}_i)}) \langle b_i \langle \hat{b}_i + 1.96(\delta_{(\hat{b}_i)})\right\}$$

$$bi = \hat{b}_i \pm (1.96).(\delta_{(\hat{b}_i)})$$

مثال حول حدود الثقة مع إختيار (Z) :

اذا علمنا أن ($\hat{b} = 8.4$) و($\delta_{(\hat{b})} = 2.2$)

بإختيار مستوى ثقة 95% نجد حدود الثقة :

$$\mathbf{b = 8.4 \pm 1.96\ (2.2)}$$

$$\mathbf{8.4 - 1.96(2.2) < b < 8.4 + 1.96(2.2)}$$

$$\mathbf{4.1 < b < 12.7}$$

وهكذا من العينة الإحصائية نستدل أن المعلمة الحقيقية المجهولة للمجتمع الإحصائي ستقع بين (4.1) و (12.7) مع إحتمال 95% .

2-6-7 حدود الثقة من إختبار (t)

إن الطريقة لبناء حدود الثقة من إختبار (t) هي مشابهة لطريقة بناء حدود الثقة من إختبار (Z) ، ولكن مع إختلاف واحد رئيس وهو في حالة إختبار (t) نأخذ في الحساب درجات الحرية (Degrees of Freedom) :

قانون (t) لإختبار ($\hat{b}_i$) مع درجات حرية (k-n) :

$$t=\frac{\hat{b}_i-b_i}{\delta_{(\hat{b}_i)}}$$

أولاً : نختار مستوى 95% مستوى ثقة أو أي مستوى ثقة آخر ، وثم نجد القيمة الجدولية أو النظرية لـ(t) المقابلة لـ($\pm t_{0.025}$) مع درجات حرية (k-n)، وهذا يتضمن إن إحتمال أن يقع (t) بين ($-t_{0.025}$) و($+t_{0.025}$) هو (0.95) مع درجات حرية (k-n) ، ويمكن كتابة الإحتمال بالطريقة الآتية :

$$P\{-t_{0.025}\langle t\langle +t_{0.025}\}=0.95$$

نعوض عن (t) بما يساوي : $\frac{\hat{b}_i-b_i}{\delta_{(\hat{b}_i)}}$

$$P\left\{-t_{0.025}\langle\frac{\hat{b}_i-b_i}{\delta_{(\hat{b}_i)}}\langle +t_{0.025}\right\}=0.95$$

$$P\left\{\hat{b}_i-t_{0.025}(\delta_{(\hat{b}_i)})\langle b_i\langle \hat{b}_i+t_{0.025}(\delta_{(\hat{b}_i)})\right\}=0.95$$

وهكذا فإن حدود الثقة بمستوى ثقة 95% بالمائة للمعلمة (b) عندما نستخدم عينة صغيرة في تقدير تلك المعلمة هي :

$$\left\{\hat{b}_i - t_{0.025}(\delta_{(\hat{b}_i)}) \langle b_i \langle \hat{b}_i + t_{0.025}(\delta_{(\hat{b}_i)})\right\}$$

مع درجات حرية k-n

أو

$$bi = \hat{b}_i \pm t_{0.025}(\delta_{(\hat{b}_i)})$$

مع درجات حرية k-n

إن معنى حدود الثقة بمستوى ثقة 95% للمعلمة (b_i) هو أن هناك إحتمال (0.95) أن تكون قيمة المعلمة الحقيقية للمجتمع الإحصائي تقع ضمن حدود الثقة ($\hat{b} \pm t_{0.025}$)(مع درجات حرية k-n) .

مثال : نفترض أننا قدرنا خط إنحدار من عينة متكونة من عشرين مشاهدة كما يلي :

$$\hat{Y} = 128.5 + 2.88X$$
$$(38.2) \quad (0.85)$$

مع درجات حرية (k-n)

18 = 2 − 20

من جدول (t) نجد أن قيمة ($t_{0.025}$) مع (18) درجة حرية هي (2.10) لذلك فإن حدود الثقة بمستوى (0.95) للمعلمات هي :

1− بالنسبة الى (b_0) : $\mathbf{b_0 = 128.5 \pm (2.10)(38.2) = 128.5 \pm 80.2}$

2− بالنسبة الى (b_1) : $\mathbf{b_1 = 2.88 \pm (2.10)(0.85) = 2.88 \pm 1.79}$

وهكذا فإن القيمة الحقيقية لمعلمة التقاطع (b_0) تقع بين (48.3) و (208.7) وبالطريقة نفسها فإن قيمة معلمة الإنحدار الحقيقي (b_1) تقع بين (1.09) و (4.67) .

الفصل الثامن

تحليل التباين

تحليل التباين

8-1 مقدمــــــــة

إن تحليل التباين Analysis of Variance (ANOVA) هــو أحــد الطرق الاحصائية التي طورت من قبــل (R.A.Fisher) لتحليــل البيانـــات التجريبية, وفي بدايتها طبقت الطريقة في تحليل التجارب الزراعية (إســتخدام أسمدة مختلفة أو بذور مختلفة) , ولكن بعد ذلك مباشرة إتسع تطبيقها ليشــمل حقول علمية أخرى في مجال البحث العلمي.

فـــي طريقـــة تحليـــل التبـــاين يمكـــن أن نقســـم التبـــاين الكلـــي (Total Variance) لمتغيـــر معـــين الـــى عناصـــر مضـــافة (Additive Component) التي ربما تعزى الى عناصر مستقلة مختلفة ، إن هذه العناصر تعبر عن أسباب أو مصادر الإنحراف أو التباين في المتغير الذي نقوم بدراســته , عنــدما تطبــق هــذه الطريقـــة علـــى البيانـــات التجريبيـــة (Experimental Data) تأخذ تصميم معين للتجربة الذي يقرر عدد العناصر الملائمة أو الاسباب للإنحراف والأهمية المنطقية لكل منها ، مثلا نفــرض ان لدينا عشرون قطعة من الارض التي نزرعها بالحنطة ونريد أن ندرس الانتـــاج لكل قطعة , نستخدم بذور مختلفة وأسمدة مختلفة وأنظمة ري مختلفة , وهكــذا فإن الإنحرافات في الانتاج ربما تعزى منطقياً الى ثلاثة عناصر :

X_1 = نوع البذور X_2 = نوع الأسمدة X_3 = نوع الري

وعند إستخدام طريقة تحليل التباين يمكن أن نقسم الإنحراف الكلي فـــي الإنتاج إلى ثلاثة عناصر مستقلة : عنصر يعود الى (X_1) وعنصر آخر يعـــود الى X_2 والعنصر الثالث الى X_3 .

من هذا التعريف لتحليل التباين يجب ان يكون واضحاً ان هذه الطريقة هي فكرياً تشبه تحليل الانحدار . ففي تحليل الانحدار فإن الهدف هـو لتقريـر العناصر التي تسبب الانحراف في قيم المتغير المعتمد .

لقد وجدنا في تحليل الانحدار ان الانحراف الكلي في قيم المتغير المعتمد ينقسم الى عنصرين : الانحراف ، الذي يوضحه خـط الانحـدار او المسـاحة المستوية ، والانحراف غير الموضح الذي تبينه نقاط شكل الانتشار حول خـط الانحـدار ، والأكثـر مـن هـذا فـإن معامـل الارتبـاط المتعـدد (Multiple Correlation Coefficient) قد عد معبـراً عـن جـزء مـن الانحراف الكلي الموضح بوساطة خط الانحدار أو المساحة المستوية ، لقد وجد أن (R^2) يساوي العناصر المسببة للإنحرافات مضافة الى بعضـها الـبعض , ولكن هناك اختلافات مهمة جداً بين طريقة تحليل التباين وطريقة الانحدار ان الاختلاف الرئيس هو ان تحليل الانحدار يزودنا بقيم رقميـة لتـأثير مختلـف العناصر التوضيحية على المتغير المعتمد ، فضلاً عن المعلومات المتعلقة بتقسيم مصادر الانحراف الكلي في (Y) في عدة عناصر مضافة بينما تحليل التبـاين يعطينا المعلومات الاخيرة حول مصادر الانحراف فقط .

ان طريقة تحليل التباين تستخدم في تحليل الانحدار لإجراء اختبـارات المعنوية أو الأهمية المختلفة ، والأكثر أهمية من تلك الاختيارات ما يأتي :

1. اختبار معنوية خط الانحدار .
2. اختبار معنوية أو أهمية التحسن في توفيق خط الانحدار الذي يحصـل من خلال ادخال متغيرات اضافية جديدة أو اضافية في الدالة .
3. اختبار التساوي بين معاملات (Coefficients) يتم الحصول عليها من عينات مختلفة .
4. اختبار العينة الاضافية من حيث أدائها في الانحدار او اختبار اسـتقرار معاملات الانحدار .

5. اختبار المقيدات المفروضة على معاملات الدالة .

8-2 : طريقة تحليل التباين بوصفها طريقة إحصائية .

إن هدف هذه الطريقة هو تقسيم الانحراف الكلي لمتغير معين (حول وسطه الحسابي) الى عدة عناصر يمكن ان تعزى الى اسباب محددة ولتبسيط التحليل سوف نفترض أن هناك عنصر نظامي واحد فقط مؤثر في التغير قيد الدراسة ، إن أي انحراف لا يسببه المتغير التوضيحي (Explanatory Variable) يفترض أن يعد انحرافا عشوائيا او إنحرافاً يحصل بالصدفة (by Chance) نتيجة لأحداث عشوائية مختلفة . ولدينا سلسلة من قيم المتغير (Y) وقيم مرافقة لها للمتغير التوضيحي (X) .

إن طريقة تحليل التباين ترتكز على قيم المتغير (Y) وتدرس انحرافاتها وقيم المتغير (X) تستخدم لتقسيم قيم (Y) في مجموعات فرعية وعينات فرعية، مثلاً مجموعة واحدة أو عينة واحدة تتضمن القيم الصغيرة للمتغير (X) .

ولكل عينة فرعية نقوم بتقدير الوسط الحسابي للمتغير (Y) بحيث نحصل على منظومة من المتوسطات الحسابية ، فإذا كان (X) (وهو أساس لتصنيف قيم Y, $Y`_s$ الى عينات فرعية) سبباً مهماً للإنحراف في (Y) (بمعنى أن X هو متغير توضيحي مهم) ، فإن الإختلاف بين المتوسطات الحسابية للعينات الفرعية سيكون كبيراً ، وهذا يبدو من خلال التشتت الكبير للمتوسطات الحسابية للعينات الفرعية لـ(Yi`s) حول الوسط الحسابي المشترك أو العام ($\overline{Y}$) . وهذا بوساطة التباين الكبير في توزيع المتوسطات الحسابية ، وعلى العكس من ذلك اذا كان (X) مصدر غير مهم من مصادر الإنحراف في (Y) فإن الإختلاف بين المتوسطات الحسابية للعينات الفرعية سوف يكون صغيراً وهذه حقيقه سوف تظهر بصيغة تباين صغير في توزيع المتوسطات الحسابية الفرعية ($\overline{Y}_i$) حول الوسط الحسابي المشترك أو العام ($\overline{Y}$) :

أ- إن أهمية المتغير (X) بوصفه سبباً للإنحراف في (Y) يحكم عليهــا من خلال الإختلاف أو الفرق (difference) بين المتوسطات الحسابية للعينات الفرعية ($\overline{Y}_i s$) التي بنيت على أساس قيم (X) .

ب- إن الفرق أو الإختلاف بين المتوسطات الحسابية تعكسه قيمة التباين في توزيع المتوسطات الحسابية للعينة .

لذلك فإن الإختلاف أو الفرق بين المتوسطات الحسابية ربمــا يــدرس ويختبر مع تقدير لتباين المجتمع الإحصــائي ، أحــدهما تقــدير ($\delta_Y{}^2$) ويــتم الحصول عليه من خلال جمع (pooling) لتباينات العينات الفرعية ، والآخــر يتم الحصول عليه من صيغة توزيع العينات أو توزيع ($\overline{Y}$) .

إن المقارنة بين أي نوعين من التباين يمكن أن تطبق بوساطة إحصاءة (F) (F statistic) وجداول (F) ، إحصاءة (F) هي نسبة أي نوعين مســتقلين من التباين الى بعضهما التي يتم الحصول عليها من بيانات العينة ، وكل تقدير يتضمن خسارة أو نقصان في درجات الحرية ، فإذا كان لدينا أي نــوعين مــن تقديرات التباين المستقل الأول مع (υ_1) من درجات الحرية (υ_2) من درجات الحرية ، فإن النسبة بين التقديرين تأخذ توزيع (F) مع درجات حرية (υ_1) و (υ_2) ، ولهذا السبب فإن (F) يدعى نســبة التبــاين (Variance Ratio) أو معدل التباين ، أما الحرف (F) فهو يرمز الى إسم (Fisher) هو إســم العــالم الذي إبتدع إحصاءة (F) .

فإذا كانت تقديرات النوعين من التباين قريبة من بعضها الــبعض فــإن النسبة بينهما سوف تقترب من قيمة الواحد ، فكلما كان الفرق بين أي نوعين من التباين كبيراً كلما كان معدل أو نسبة (F) كبيراً ، وهكذا وبعامة فإن القيم العالية لــ(F) تعبر عن إن الفرق بين النوعين من التباين ذا أهمية كبيــرة أو رفــض

فرضية العدم (Null Hypothesis) التي تفترض أن ليس هنـــاك إخـــتلاف أو فرق مهم بين النوعين من التباين .

3-8 إختبار الفرق بين المتوسطات الحسابية

نفترض ثلاثة أنواع من البترول تستخدم لتشغيل السيارة :

النوع (A) الذي يتصف بـــ(90) أوكتان (Octane)

النوع (B) الذي يتصف بـــ(95) أوكتان

النوع (C) الذي يتصف بـــ(100) أوكتان

وهنا نرغب في أن نعرف في ما اذا كانت هذه النـــواع الثلاثـــة مـــن البترول تعطي الإستهلاك نفسه من البترول لكل ميل أم لا ؟ بمعنى آخر نريد أن نقارن أداء الإستهلاك للإنواع الثلاثة من البترول ، لنفترض أننا نســتخدم كـــل نوع لمدة (10) أيام ونقيس عدد الأميال المقطوعة لكل غـــالون مـــن البتـــرول وهكذا نحصل على ثلاث عينات بحجم (10) مشاهدات لكل نوع من البترول ، وهذه المشاهدات مدونة في الجدول (8-1) بصيغة عدد الأميال لكل غالون من البترول .

جدول (8-1)

العينة (1) نوع البترول (A) **N1 = 10**	العينة (2) نوع البترول (B) **N2 = 10**	العينة (3) نوع البترول (C) **N3 = 10**	المجموع **N=n1+n2+n3**
32	**35**	**44**	
30	**38**	**46**	
35	**37**	**47**	
33	**40**	**47**	
35	**41**	**46**	
34	**35**	**43**	
29	**37**	**47**	
32	**41**	**45**	
36	**36**	**48**	
34	**40**	**47**	
$\sum Y_{1i} = 330$	$\sum Y_{2i} = 380$	$\sum Y_{3i} = 460$	
$\bar{Y} = \frac{\sum Y_{1i}}{n_1} = 33$	$\bar{Y}_2 = \frac{\sum Y_{2i}}{n_2} = 38$	$\bar{Y}_3 = \frac{\sum Y_{3i}}{n_3} = 46$	
$\delta^2_1 = \frac{\sum_i^{n_1}(Y_{1i}-\bar{Y}_1)^2}{n_1}$	$\delta^2_2 = \frac{\sum_i^{n_2}(Y_{2i}-\bar{Y}_2)^2}{n_2}$	$\delta^2_3 = \frac{\sum_i^{n_3}(Y_{3i}-\bar{Y}_3)^2}{n_3}$	
$= \frac{46}{10}$	$= \frac{50}{10}$	$= \frac{22}{10}$	$\bar{Y} = \frac{\sum_i \sum_i Y_{1i}}{N}$
$\delta^2_1 = 4.6$	$\delta^2_2 = 5.0$	$\delta^2_3 = 2.2$	**= 39**

المجموع **N**
32
30
35
33
35
34
29
32
36
34
35
38
37
40
41
35
37
41
36
40
44
46
47
47
46
43
47
45
48
47
$\sum_i \sum_i Y_{ii} =$
= 1170

يمكن أن تفسر البيانات التي وضعت في الجدول (8−1) بوصفها ثلاثة عينـــات عشـــوائية بحجـــوم ($n_1=n_2=n_3=10$) مـــع متوســـطات حســـابية ($\overline{Y}_1=33$ ، $\overline{Y}_2=38$ ، $\overline{Y}_3=46$) ميل/غالون من البترول .

مشكلتنا هنا هي أن نتأكد هل الإختلاف أو الفرق بين هذه المتوســـطات الحسابية مهم إحصائياً أو يمكن أن يعزى الى الصـــدفة (Chance) يجـــب أن نفترض أن العينات مأخوذة من ثلاثة مجتمعات إحصائية تتميز بأن لها توزيـــع طبيعي أو تقريباً طبيعي مع متوسطات حسابية (M_3,M_2,M_1) على التوالي مع إنحراف معياري (δ) متساو ، إن هذا الإفتراض يتضمن أنه على الرغم مـــن الإختلاف في محتوى البترول من الأوكتان من الأنواع الثلاثة من البترول ربما تؤثر على معدل إستهلاك البترول ، فإنه سوف لن يؤثر التشتت (dispersion) أو التباين في الأميال المقطوعة حول المتوسطات الحسابية ، بعبارة أخـــرى إذا أخذنا عدداً كبيراً من المشاهدات لكل نوع من البترول فإن التوزيعـــات الثلاثـــة التي سوف نحصل عليها سوف تكون قريبة من المنحنيات الطبيعيـــة ، حيـــث تتصف التوزيعات بالإنحراف المعياري نفسه (δ) ، نريد أن نعرف في مـــا إذا كان هناك ثمة إختلاف أو فرق مهـــم إحصـــائياً بـــين المتوســـطات الحســـابية (μ_3,μ_2,μ_1) للمجتمعات الإحصائية الثلاثة :

نريد أن نختبر فرضية العدم (Null Hypothesis)

$$H_0 : \mu_2 = \mu_3 = \mu_1$$

مقابل الفرضية البديلة

$$H_1 : \mu_j \quad \text{غير متساوية}$$

فإذا كانت المتوسطات الحسابية الثلاثة متساوية وهذا يعنـــي اذا كانـــت فرضية العدم صحيحة فإن المجتمعات الإحصائية الثلاثة ربمـــا تعـــد بوصـــفها مجتمعاً إحصائياً كبيراً واحداً مع وسط حسابي :

$(\mu_1=\mu_2=\mu_3=)\mu$ مع إنحراف معياري (δ) وهذا يعني :

$$Y \sim N(\mu, \delta\ \ )$$

والعينات الثلاثة ربما تعد بوصفها عينات أخذت من مجتمع إحصائي واحد كبير ، وربما تكتب التوزيعات الآتية للمتوسطات الحسابية للعينات : $(\overline{Y}_1\ \ \overline{Y}_2\ \ \overline{Y}_3)$

$$\overline{Y}_1 \sim N(\mu, \delta^2_{\overline{Y}_1}) \sim N(\mu, \frac{\delta^2}{n_1})$$

$$\overline{Y}_2 \sim N(\mu, \delta^2_{\overline{Y}_2}) \sim N(\mu, \frac{\delta^2}{n_2})$$

$$\overline{Y}_3 \sim N(\mu, \delta^2_{\overline{Y}_3}) \sim N(\mu, \frac{\delta^2}{n_3})$$

قلنا ضمن فرضية العدم $(\mu_1=\mu_2=\mu_3)$ وربما نعد المجتمعات الإحصائية الثلاثة تشكل مجتمع إحصائي كبير :

$$Y \sim N(\mu, \delta^2)$$

إن تقديراً للوسط الحسابي العام أو المشترك (μ) ربما يحسب من خلال توسيع العينة $(n_1=n_2=n_3=30)$ ومن البيانات الموجودة في الجدول $(8-1)$ نحصل على :

$$\hat{\mu} = \frac{\sum Y_i}{N} = \frac{\sum_{j}^{k}\sum_{i}^{nk} Y_{ji}}{N} = \frac{1170}{30} = 39 = \overline{Y}$$

إن تقديراً لتباين المجتمع الإحصائي (δ^2) ربما تحصل بطريقتين :

أولاً : يمكن أن نحصل على تقدير غير متحيز لتباين المجتمع الإحصائي من الصيغة الآتية :

$$\hat{\delta}^2 = \frac{\sum_{j}^{k} n_j(\overline{Y}_j - \overline{Y})^2}{k-1} \qquad \ldots\ldots\ldots\ldots\ldots (1-8)$$

عندما تكون (k) عدد العينات المأخوذة .

البرهان : إن هذه الصيغة إشتقت من العلاقة بين تباين المجتمع الإحصائي (δ^2) وتباين توزيع العينات :

$$\delta^2_{\overline{Y}_j} = \frac{\delta^2}{n_j} \cdots or \cdots \delta^2 = \delta^2_{\overline{Y}_j}.n_j$$

في مثالنا لدينا ثلاثة عينات ومن كل منها يمكن أن نحصل على تقدير مستقل لـ(δ^2) :

$$\hat{\delta}^2_1 = n_1.\delta^2_{\overline{Y}_1} = n_1(\overline{Y}_1 - \overline{Y})^2 \; 10(33-39)^2=360$$

$$\hat{\delta}^2_2 = n_2.\delta^2_{\overline{Y}_2} = n_2(\overline{Y}_2 - \overline{Y})^2 \; 10)38-39)^2=10$$

$$\hat{\delta}^2_3 = n_3.\delta^2_{\overline{Y}_3} = n_3(\overline{Y}_3 - \overline{Y})^2 \; 10(46-39)^2=490$$

عندما يكون ($\overline{Y}$) الوسط الحسابي المشترك أو العام (pooled) نأخذ المعدل الموزون (Weighted Average) لهذه التقديرات نحصل على :

$$\hat{\delta}^2 = \frac{1}{3}\sum_{j}^{k} n_j(\overline{Y}_j - \overline{Y})^2$$

من أجل الحصول على تقدير غير متحيز نستخدم درجات حرية (3−1=2) ، أو على نحو عام (1-k) اذا كان لدينا (k) من العينات ، وهكذا فإن التقدير الأول لتباين المجتمع الإحصائي يصبح :

$$\hat{\delta}^2 = \frac{\sum_{j=1}^{k} n_j(\overline{Y}_j - \overline{Y})^2}{k-1}$$

يجب أن يكون واضحاً الآن أن هذا التقدير لتباين المجتمع الإحصائي تم الحصول عليه من الفروقات بين المتوسطات الحسابية ($\overline{Y}_j$) والوسط الحسابي المشترك ($\overline{Y}$) للمجتمع الإحصائي .

وكما رأينا فإن المتوسطات الحسابية للعينات الإحصائية هي تقديرات غير متحيزة للمتوسطات ($\mu_3=\mu_2=\mu_1$) ، إن فرضية العدم كانت ($\mu_3=\mu_2=\mu_1$) $=\mu^1$ لذلك اذا كانت هذه الفرضية صحيحة فإن المتوسطات

الحسابية للعينات ($\overline{Y}_3 \;\; \overline{Y}_2 \;\; \overline{Y}_1$) يجب أن لا تختلف عن بعضها الـــبعض ولا تختلف عن الوسط الحسابي للمجتمع الإحصائي ($\overline{Y}$) ، إن هذا يتضمن اذا كانت تلك الفرضية غير صحيحة فإننا نتوقـــع إن المتوســـطات الحســـابية للعينـــات ($\overline{Y}_3 \;\; \overline{Y}_2 \;\; \overline{Y}_1$) ، يجب أن تختلف عن بعضها البعض بشكل كبير وكذلك تختلف عن الوسط الحسابي للمجتمع الإحصائي ($\overline{Y}$) : إن الفرق بين هذه المتوسطات الحسابية سيكون أكبر مما نعزيه الى الصدفة ، وهذا بدوره يتضـــمن أن تقـــدير ($\hat{\delta}^2$) لتباين المجتمع الإحصائي سوف يكون كبيراً اذا كانت فرضية العدم غير صحيحة ، وذلك لأن ($\hat{\delta}^2$) قد تم إحتسابه من الفروقات $(\overline{Y}_j - \overline{Y})^2$ وهكذا فإن التقدير ($\hat{\delta}^2$) هو عنصر حاسم في إختبار الفرق بـــين المتوســـطات الحســـابية لأحجام معينة من العينات (Samples) . من خلال طريقة تقديره يتبـــين أنـــه يعكس الإنحرافات (Variations) بين المتوسطات الحسابية للعينـــات ويـــدعى الإنحراف بين (Variation between) . ومن أجل إجراء إختبارنا يكفـــي أن نقارن التقدير مع تباين المجتمع الإحصائي الحقيقي($\hat{\delta}^2$) ونرفض فرضية العدم اذا كان الإنحراف أو التباعد (divergence) بين ($\hat{\delta}^2$)(δ^2) كبيراً ، ولكن في مثالنا وكما هي الحال في أغلب المشكلات الموجودة في عالم الواقع ، إن (δ^2) الحقيقي غير معروف مما يوجب علينا أن نحصل على قيمـــة أخـــرى مقـــدرة مستقلة من بيانات عينة إحصائية .

ثانياً : إن تقديراً لتباين المجتمع الإحصائي ($\hat{\delta}^2$) يمكن أن نحصل عليه من خلال عملية تباينات العينات الإحصائية ، والصيغة أو القانون الملائم هو :

$$\hat{\hat{\delta}}^2 = \frac{n_1 s^2 + n_2 s^2 + \dots\dots\dots + n_k s^2{}_k}{(n_1 + n_2 + \dots\dots + n_k) - k}$$

عندما يكون : (s_j) = تباينات العينات الإحصائية

(n_j) = حجوم العينات

يجب أن نلاحظ أن :

$$\left.\begin{aligned} n_1\delta^2{}_1 &= n_1 \frac{\sum (Y_{1i}-\overline{Y})^2}{n_1} \sum (Y_i-\overline{Y}_1)^2 \\ n_k\delta^2{}_k &= n_k \frac{\sum_{j}^{nk} Y_{ki}-Y_k)^2}{n_k} = \sum (Y_{ki}-\overline{Y}_k)^2 \end{aligned}\right\} \ldots\ldots(3-8)$$

وهكذا فإن $\left(n_1\delta^2{}_1+n_2\delta^2{}_2+\ldots\ldots\ldots+n_k\delta^2{}_k\right)$ تعطي المجموع الكلـــي لتربيـــع الإنحرافات لكل العينات (k) ، إن التباين المجمـــوعي (pooled-Variance) وصيغته يمكن أن يعد بوصفه عملية ربط أو جمع كل العينات في عينة واحـــدة كبيرة وتقدير تباين المجتمع الإحصائي ، نعوض عن (8−3) في (8−2) نحصل على :

$$\hat{\hat{\delta}}^2 = \frac{\sum_{i}^{n_1}(\overline{Y}_{1i}-\overline{Y}_1)^2+\sum_{i}^{n_2}(Y_{2i}-\overline{Y}_2)^2+\ldots\ldots+\sum_{i}^{n_k}Y_{ki}-\overline{Y}_k)^2}{N-k}$$

عندما يكون $n_k\ldots\ldots\ldots\ldots+n_2+n_1=N$

وبإستخدام إشارة أو علامة التجميع المزدوج نحصل على :

$$\hat{\hat{\delta}}^2 = \frac{\sum_{j}^{k}\sum_{i}^{n_k}(\overline{Y}_{ji}-\overline{Y}_j)^2}{N-k} \quad \ldots\ldots\ldots\ldots\ldots(4-8)$$

إن هذا التقدير لتباين المجتمع الإحصائي تم الحصول عليـــه أو يمكـــن الحصول عليه من تباينـــات العينـــات (Sample Variances) التـــي تعكـــس الإنحراف ضمن كل عينة . كما أن تباينات العينات لا تعتمد على فرضية العدم وتتأثر تلك التباينات بالفروقات (differences) بـــين المتوســـطات الحســـابية للعينات $\left(\overline{Y}_3 \;\; \overline{Y}_2 \;\; \overline{Y}_1\right)$. بعبارة أخرى نقول حتـــى لـــو كانـــت المتوســـطات الحسابية مختلفة على نحو كبير ، بحيث يكون هناك ثلاثة مجتمعات إحصـــائية

لكل واحد منها وسطه الحسابي ($\mu_3 = \mu_2 = \mu_1$) ولكن كل تلك المجتمعات الإحصائية سوف يكون لها بالإفتراض التباين نفسه (δ^2) للمجتمعات الإحصائية المجمعة (pooled) ($\hat{\delta}^2$) الذي يعتمد على الإنحراف ضمن (Variation within) قيم العينة ($Y_{i's}$) لكل عينة) والذي يدعى الإنحراف ضمن (within Variation) ، نلاحظ الآن أن الإنحراف في قيم (Y_i) في كل عينة يعبر عن إنحرافات بالصدفة ، لذلك فإن التقدير ($\hat{\delta}^2$) ربما يعد مقياساً للإنحراف في قيم (Y_i) ذلك الإنحراف الذي يمكن أن يعزى الى الصدفة .

الآن لدينا تقديرين غير متحيزين (Unbaised) لتباين المجتمع الإحصائي (δ^2) :

التقدير (1) يعكس الإنحراف بين المتوسطات الحسابية للعينات ويعتمد على شرعية أو صحة فرضية العدم .

التقدير (2) يعكس الإنحراف في قيم (Y_i) ضمن العينات ومستقل عن فرضية العدم .

إن هذين التقديرين مستقلين ، لذلك فإن النسبة بينهما يكون لهما توزيع (F distribution)(F) مع درجات حرية كما يأتي ($k-1 = \upsilon_1$) و ($N-k = \upsilon_2$) .

$$F^* = \frac{\left[\sum_{j=1}^{k} n_j(\overline{Y}_j - \overline{Y})^2\right] / (k-1)}{\left[\sum_{j=1}^{k}\sum_{i=1}^{nj}(Y_{ji} - \overline{Y}_j)^2\right] / (N-k)}$$

عندما يكون :

Nj= حجم العينة (jth Sample)

$N = \sum_{j=1}^{k} n_j$ = حجم العينة المكبرة الناتجة عن تجميع العينات

K = عدد العينات

وهكذا يمكن وضع معدل أو نسبة التباين (Variance Ratio) بصيغة منتظمة :

$$F^* = \frac{\text{التباين المقدر من الإنحراف بين المتوسطات الحسابية للعينات}}{\text{التباين المقدر من الإنحراف في قيم } (Y_i) \text{ للعينات}}$$

عندما لا تكون المتوسطات الحسابية ($\mu_3=\mu_2=\mu_1$) مساوية للتباين المقدر من الفروقات بين المتوسطات الحسابية فإنها تكون كبيرة ولذلك فإن معدل التباين (F^*) سوف يصبح كبيراً ، فإذا كانت فرضية العدم صحيحة فإن نسبة التباين المشاهد سوف تقترب من قيمة الواحد :

إن الفرق المشاهد في المتوسطات الحسابية ($\overline{Y}_1$ $\overline{Y}_2$ $\overline{Y}_3$) في هذه الحالة غير أساسي أو غير مهم ، وربما يعزى الى الصدفة ، وهكذا فإن التقدير الذي يبدو في البسط (Numerator) من النسبة (F^*) سيكون في الحقيقة تقدير لتباين المجتمع الإحصائي المجهول نفسه الذي يقدر في المقام (denominator) في النسبة(F^*) .

إن نسبة التباين المشاهد (F^*) تقارن مع القيمة النظرية (The Theoretical Value of F) لـ(F) (مع مستوى معنوية مختار مثلاً 5%) وتلك القيمة النظرية نحصل عليها من جدول(F) مع :

($k-1= \upsilon_1$) و ($N-k = \upsilon_2$) درجات حرية .

إن القيمة النظرية (القيمة الحاسمة) لـ(F) هي قيمة (F) التي تعرف لنا المنطقة الحاسمة أو الحرجة للإختبار بمستوى مختار من المعنوية ، اذا كانت (F^*) أكبر من (F) النظرية نرفض فرضية العدم ، وهذا يعني أننا نقبل إن الفرق بين المتوسطات الحسابية كبير ومهم ، ومن هذه الشواهد ربما نفهم أو

نستنتج أن المجتمعات الإحصائية التي تم سحب العينات الإحصائية منها تختلف عن بعضها البعض .

واذا كانت (F^*) أصغر من (F) النظرية نقبل فرضية العدم وهذا يعني أننا نقبل أن المتوسطات الحسابية للعينات لا تختلف على نحو كبير أو مهم .

وفي هذه الحالة ربما نقول أن بيانات العينة تزودنا بدليل بأنه ليس هناك إختلاف مهم أو كبير بين المتوسطات الحسابية للمجتمعات الإحصائية التي تم سحب العينات منها .

في مثالنا لدينا النتائج الآتية :

(1) تباين مقدر بين المتوسطات الحسابية للعينات :

$$\hat{\delta}^2 = \frac{\sum_{j=1}^{k} n_j(\overline{Y}_j - \overline{Y})^2}{k-1}$$

$$= \frac{n_1(\overline{Y}_1 - \overline{Y})^2 + n_2(\overline{Y}_2 - \overline{Y})^2 + n_3(\overline{Y}_3 - \overline{Y})^2}{3-1}$$

$$= \frac{10(33-39)^2 + 10(38-39)^2 + 10(46-39)^2}{2}$$

$$= 430$$

(2) تقدير التباين ضمن القيم (within Variation) :

$$\hat{\hat{\delta}}^2 = \frac{\sum_{j}^{k}\sum_{i}^{n_k}(\overline{Y}_{ji} - \overline{Y})^2}{N-k}$$

$$= \frac{\sum_{1}^{10}(\overline{Y}_{1j} - \overline{Y}_1)^2 + \sum_{1}^{10}(Y_{2j} - \overline{Y}_2)^2 + \sum_{1}^{10}(Y_{3j} - \overline{Y}_3)^2}{30-3}$$

$$= \frac{46+50+22}{27} = \frac{118}{27}$$

$$\approx 4.37$$

(3) نسبة التباين المشاهد(The Observed Variance Ratio) :

$$F^* = \frac{\left[\sum_{j=1}^{k} n_j(\bar{Y}_j - \bar{Y})^2\right]/(k-1)}{\left[\sum_{j=1}^{k}\sum_{i=1}^{nj}(Y_{ji} - \bar{Y})^2\right]/(N-k)}$$

$$F^* = \frac{\hat{\delta}^2}{\hat{\hat{\delta}}^2} = \frac{430}{4.37}$$

$$F^* = 98.39 \approx$$

(4) القيمة النظرية لــ(F^*) بمستوى معنوية (5%) مع درجات حريــة ($k-1= \upsilon_1$)و(υ_2 = N-k = 27) وجدت من جدول (F) كما يأتي :

$$F_{0.05} = 3.37$$

(5) لإن (F^*) أكبر من $F_{0.05} = 3.37$ نرفض فرضية العدم ، بمعنــى أننا نقبل أن هناك ثمة فارق مهم في معدل الأميال التي تحصل مــن الأنــواع الثلاثة للبترول ، إن الإختبار المذكور آنفاً ربما يفحص بطريقة أخرى ســوف تنظم طريقة تحليل التباين ، يمكن أن نحصل على تقدير ثالث لتباين المجتمــع الإحصائي (δ^2) بإستخدام عينة موسعة أو مكبرة تشكل من العينــات الثلاثــة الفرعية والتقدير غير المتحيز سيكون :

$$^*\delta^2 = \frac{\sum_{i}^{N}(Y_i - \bar{Y})^2}{N-1} = \frac{\sum_{j=1}^{k}\sum_{i=1}^{nj}(Y_{ji} - \bar{Y})^2}{N-1}$$

عندما يكون N-1= درجات الحرية للتباين المقدر (δ^2*) فــإذا أخــذنا البسط للتقديرات الثلاثة لتباين المجتمع الإحصائي ($^*\delta^2, \hat{\delta}^2, \hat{\hat{\delta}}^2$) يمكننا أن نضع العلاقة الآتية بين تلك الصيغ :

$$\sum_{j=1}^{k}\sum_{i=1}^{nj}(Y_{ji}-\overline{Y})^2=\sum_{j=1}^{k}n_j(\overline{Y}_j-\overline{Y})^2+\sum_{j=1}^{k}\sum_{i=1}^{nj}(Y_{ji}-\overline{Y}_j)^2$$

(المجموع الكلي لمربع الإنحرافات) = (مجموع مربعات الإنحرافات بين المجموعات) + (مجموع مربعات الإنحرافات ضمن المجموعات)

وهذا يعني :

(الإنحراف الكلي في (Y)) = (الإنحرافات بين المتوسطات) + (الإنحرافات ضمن العينات)

البرهان : نبدأ من الصيغة على الجهة اليسرى ونشكل المطابقة :

$$\left(Y_{ji}-\overline{Y}\right)=\left(Y_{ji}-\overline{Y}\right)+\left(\overline{Y}_j-\overline{Y}_j\right)$$

أو

$$\left(Y_{ji}-\overline{Y}\right)=\left(Y_{ji}-\overline{Y}_j\right)+\left(\overline{Y}_j-\overline{Y}\right)$$

نقوم بتربيع الطرفين :

$$\left(Y_{ji}-\overline{Y}\right)^2=\left(Y_{ji}-\overline{Y}\right)^2+\left(\overline{Y}_j-\overline{Y}_j\right)^2+2\left(Y_{ji}-\overline{Y}\right)=\left(Y_{ji}-\overline{Y}_j\right)+\left(\overline{Y}_j-\overline{Y}\right)$$

نجمع كل القيم فنجد :

$$\sum_{j=1}^{k}\sum_{i=1}^{nj}(Y_{ji}-\overline{Y})^2=\sum_{j}\sum_{i}(\overline{Y}_j-\overline{Y})^2+\sum_{j}\sum_{i}(\overline{Y}_j-\overline{Y}_j)^2+2\sum_{j}\sum_{i}(Y_{ji}-\overline{Y}_j)(\overline{Y}_j-\overline{Y})$$

إن الجزء الأخير من هذه المتطابقة يساوي صفراً ، وطالما أن :

$$2\sum_{j}\sum_{i}\left(Y_{ji}-\overline{Y}_j\right)\left(\overline{Y}_j-\overline{Y}\right)=2\sum_{j}\left[\left(\overline{Y}_j-\overline{Y}\right)\sum\left(Y_{ji}-\overline{Y}_j\right)\right]$$

واذا علمنا أن ($\sum_{i}\left(Y_{ji}-\overline{Y}_j\right)=0$) وذلك لأن هذه هي مجموع الإنحرافات ضمن المجموعات (groups) أو العينات لذلك :

$$\sum_{j}^{k}\sum_{i}^{nj}(Y_{ji}-\overline{Y})^2=\sum_{j}^{k}\sum_{i}^{nj}(Y_{ji}-\overline{Y}_j)^2+\sum_{j}^{k}n_j(\overline{Y}_j-\overline{Y})^2$$

الكلي = إنحرافات ضمن + إنحرافات بين

إن هذا التعبير أو الصيغة يبين كيف أن المجموع الكلي لمربعات في (Y)(في المجموعات مع بعض) ينقسم الى جزئين :

الجزء الأول من الإنحرافات في (Y) يعود الى الفرق بين المتوسطات الحسابية الجـــزء الثـــاني مـــن الإنحرافـــات فـــي (Y) يعـــود الـــى الصـــدفة (مثلاً المطر أو العوامل النفسية) .

لاحظ أن هذا التقسيم للإنحراف الكلي الى أجزاء مضافة يبقى صـــحيحاً بغض النظر عن قبول فرضية العدم أم لا ، وفي مثالنا فإن الإنحراف الكلي في قيم (Y) حول الوسط الحسابي المشترك ($\overline{Y}$ =39) هو :

$$\sum_{j}^{3}\sum_{i}^{10}\left(Y_{ji}-\overline{Y}\right)^{2}=978$$

ومربعات الإنحرافات بين المتوسطات الحسابية :

$$\sum_{j=1}^{3}n_{j}(\overline{Y}_{j}-\overline{Y})^{2}=860$$

مربعات الإنحرافات ضمن المجموعات أو العينات

$$\sum_{j}^{3}\sum_{i}^{10}\left(Y_{ji}-\overline{Y}_{j}\right)^{2}=118$$

واضحاً :

ضمن + بين = المجموع

978 = 860 + 118

وهذا التقسيم للتباين الكلي يبقى صحيحاً على الرغم من إثبـــات فشـــل فرضية العدم ، وهنا يمكننا أن نضع العلاقة بين درجات الحرية في كل واحـــد من التقديرات الثلاثة لتباين المجتمع الإحصائي :

أ- إن درجات الحرية للتباين الكلي ($\delta^{2}*$) هي (N-1) .

ب- إن درجات الحرية للتباين المقدر على أساس الفرق بين المتوسطات الحسابية ($\hat{\delta}^{2}$) هي (K-1) .

ج- إن درجات الحرية للتباين المقدر على أساس الإنحرافات ضمن المجموعات (العينات) $(\hat{\hat{\delta}}^2)$ هي (N-K) .

الآن من السهولة أن نضع الإنحراف الكلي بالصيغة الآتية :

$$(N-1) = (N-K) + (K-1)$$

بين + ضمن = الكلي

مع المعلومات حول تقسيم الإنحرافات الكلية (الإنحراف الكلي فـــي Y) ودرجات الحرية المختلفة يمكننا أن نشكل جدول تحليل التباين :

مصدر الإنحراف (1)	مجموع المربعات (2)	درجات الحرية (3)	متوسط المربعات (4) = (2) : (3)	F (5)
بين المتوسطات والوسط الحسابي العام ضمن العينات	$\sum_{j}^{k} n_j (\bar{Y}_j - \bar{Y})^2$ $\sum_{j}^{k} \sum_{i}^{n_j} \left(Y_{ji} - \bar{Y}_j \right)^2$	$\upsilon_1 = k-1$ $\upsilon_2 = N-k$	$\frac{\sum_{j} n_j (\bar{Y}_j - \bar{Y})^2}{k-1}$ $\frac{\sum_{j} \sum_{i} (\bar{Y}_{ji} - \bar{Y})^2}{N-k}$	$F^* = \frac{\sum_{j} n_j \left(\bar{Y}_j - \bar{Y} \right)^2 /(k-1)}{\sum_{j} \sum_{i} \left(Y_{ji} - \bar{Y}_j \right)^2 /(N-K)}$
الإنحراف الكلي	$\sum_{j}^{k} \sum_{i}^{n_j} \left(Y_{ji} - \bar{Y} \right)^2$	(N-1)		من الجدول (F) $\upsilon_1 = k-1$ $\upsilon_2 = N-k$

ملحوظة : k هو عدد الهينات المأخوذة من المجتمع الإحصائي

$$F^* = \frac{\text{بين}}{\text{ضمن}}$$

إن نسبة (F^*) نسبة التباين المشاهد تشكل من خلال تقسـيم متوسـطي مربع الإنحراف أو الأخطاء التي تبدو في العمود الرابع من جدول تحليل التباين في مثالنا الرقمي فإن جـدول تحليـل التبـاين يبـدو فـي الجـدول (3-8) الآتي :

جدول (8-3)

مصدر الإنحراف	مجموع مربعات الإنحرافات	درجات الحرية	متوسط المربعات	F*
بين ضمن	**860** **118**	**(3−1)=2** **(30−3)=27**	$\frac{860}{2} = 430$ $\frac{118}{27} = 4.37$	$F^* = \frac{430}{4.37}$ **= 98.4**
المجموع	**978**	**(30−1)=29**		**من جدول F** $F_{0.025}=3.37$ **مع درجات حرية** $\upsilon_1=2$ $\upsilon_2=27$

8-4 إختبار (F) لمعنوية خط الإنحدار

في هذا المجال نحاول إختبار نماذج تحتوي على متغيرات توضيحية عدة، إن الإختبار يهدف الى تبيان في ما اذا كانت المتغيرات التوضيحية تؤثر فعلاً وعلى نحو مهم على المتغير المعتمد ، ويمكن أن يأخذ إختبار معنوية خط الإنحدار الشكل الآتي الذي يتضمن فرضية العدم :

$$H_0 : b_1 = b_2 = \dots\dots\dots = b_k = 0$$

ضد أو مقابل الفرضية البديلة :

H_1 : إن كل b_i لاتساوي صفر

اذا كانت فرضية العدم صحيحة أو حقيقية (True) بمعنى اذا كانت كل المعلمات الحقيقية مساوية للصفر ، أي ليس هناك هلاقة خطية بين المتغير (Y) والمتغيرات التوضيحية ، ولكن إختبار المعنوية الكلية للإنحدار يمكن أن يجري بإستخدام جدول تحليل التباين .

نحسب إنحدار (Y) على كل (X) مع بعض ونقدر :

أ- المجموع الكلي لمربعات الإنحرافات لـــ$((Y_i - \overline{Y}))$ $\left(\sum (Y_i - \overline{Y})\right)$

ب- مجموع مربعات الإنحرافات الموضحة من قبل المتغيرات التوضيحية مـــع بعض $\left(\sum (\hat{Y} - \overline{Y})\right)$

ج- مجموع الإنحرافات المتبقية وهي $\left(\sum \ell_1^2\right)$

من هذه الصيغ (terms) نستطيع أ، نضع التعبير الآتي :

$$\sum Y^2 = \sum \hat{Y}^2 + \sum \ell^2$$

نحصل على درجات الحرية لكل صيغة في المتطابقة :

درجات الحرية للصيغة $\left(\sum \hat{Y}\right)$ هي (k-1) عندما يكون k(=k+1) هو العدد الكلي للمعلمات ($b's$) ومن ضمنها معلمة التقاطع .

درجات الحرية للصيغة $\left(\sum \ell^2\right)$ هي (N-K) عندما يكون (N) حجـــم العينة الإحصائية .

وأخيراً فإن درجات الحرية لمجموع المربعات هو :

$$(K-1)+(N-K) = N-1$$

وضمن هذه المعلومات ربما نحسب نسبة (F^*) بالصيغة الآتية :

$$F^* = \frac{\sum \hat{y}^2 /(K-1)}{\sum \ell^2 /(N-K)}$$

ونتيجة هذه الصيغة تقارن مع القيمة الجدولية لـــــ(F) (فـــي مســـتوى مختار من المعنوية) مع درجات حرية :

$$(N - k = \upsilon_2) \qquad (k - 1 = \upsilon_1)$$

اذا كانت (F^*) أكبر من (F) نرفض فرضية العدم بمعنى أننا نقبـــل أن الإنحدار ذا معنوية إحصائية عالية ، أي أنه ليس كل المعلمات تساوي صفراً ، و اذا كانت (F^*) أصغر من (F) نقبل فرضية العدم بمعنى أننا نقبـــل فرضـــية

العدم أن الإنحدار ليس ذا معنوية إحصائية عالية ، يمكن أن نلخص المعلومات المذكورة آنفاً بوصفها في جدول تحليل التباين :

مصدر الإنحراف	مجموع المربعات SS	درجات الحرية	متوسط مربعات الأخطاء	F*
$X_1,X_2,X_4 \ldots\ldots\ldots X_K$	$\sum \hat{y}^2$ $\sum \ell^2$	υ_1 =k-1 υ_2 =N-k	$\frac{\sum \hat{y}^2}{k-1}$ $\frac{\sum \ell^2}{N-k}$	$F^* = \frac{\sum \hat{y}^2 /(K-1)}{\sum \ell^2 /(N-K)}$ $\frac{R^2 /(k-1)}{(1-R^2)/(N-k)}$
المجموع	$\sum y^2$	N - 1		من الجداول مع درجات معنوية F υ_1 =k-1 υ_2 =N-k

يمكن أن نبين أن نسبة (F) لإختبار المعنوية الكلية للإنحدار يمكـــن أن تقلص الى

$$F = \frac{R^2 /(K-1)}{(1-R^2)/(N-K)}$$

عندما يكون :

K = عدد المعلمات (b'_s) ومن ضمنها (b_0) معلمة التقاطع .

N = عدد المشاهدات في العينة

البرهان :

معلوم لدينا أن :

$$F^* = \frac{\sum \hat{y}^2 /(K-1)}{\sum \ell^2 /(N-K)}$$

نعيد كتابة الصيغة أعلاه بالصيغة الآتية :

$$F^* = \frac{\sum \hat{y}^2}{\sum \ell^2} \cdot \frac{N-K}{K-1}$$

نقسم البسط والمقام على $(\sum y^2)$ فنحصل على :

$$F^* = \frac{\frac{\sum \hat{y}^2}{\sum y^2}}{\frac{\sum \ell^2}{\sum y^2}} \cdot \frac{N-K}{K-1}$$

ومعلوم أيضاً أن :

$$\frac{\sum \hat{y}^2}{\sum y^2} = R^2{}_{Y.X_1.X_2....X_k}$$

وإن :

$$\frac{\sum \ell^2}{\sum y^2} = 1 - R^2{}_{Y.X_1.X_2....X_k}$$

نعوض في (F^*) نجد أن :

$$F^* = \frac{R^2{}_{Y.X_1,X_2,...X_k}}{1 - R^2{}_{Y.X_1,X_2,...X_k}} \cdot \frac{N-K}{K-1} = \frac{R^2/(K-1)}{(1-R^2)/(N-K)}$$

يمكن ان نقول اذا وجد أن الإنحدار ليس له معنوية عالية أي اذا كانت المتغيرات التوضيحية لاتوضح أو لاتفسر أية نسبة من الإنحرافات في المتغر المعتمد ، فإننا نتوقع أن قيمة البسط في صيغة (F^*) سوف تكون صغيرة جداً ، بعبارة أخرى أن (F^*) سوف يقترب من الصفر ، إن القيمة الكبيرة لـ (F^*) تعني أن علاقة ذات معنوية إحصائية عالية ومهمة يعرضها الإنحدار بين (Y) و (X_S) .

الفصل التاسع

مشكلات الإنحدار

مشكلة الإرتباط الخطي المتعدد

مشكلات الإنحدار

مشكلة الإرتباط الخطي المتعدد (Multicollinearity)

9-1 مقدمـــــــة

من الشروط الحاسمة لتطبيق طريقة المربعات الصغرى هو أن المتغيرات الخطية لا ترتبط ببعضها خطياً وبشكل تام أي أن ($Y_{X_iX_j} \neq 1$) ، إن التعبير أو المصطلح (Multicollinearity) يستخدم لتعبير عن وجود علاقات خطية أو علاقات قريبة من الخطية بين المتغيرات التوضيحية (Explanatory Variables) . فإذا إرتبطت المتغيرات التوضيحية ببعضها خطياً وبشكل تام أي إذا كان معامل الإرتباط لتلك المتغيرات يساوي واحد ، فإن المعلمات المقدرة تصبح غير محدودة أو غير نهائية : بمعنى أنه من المستحيل الحصول على قيم رقمية لكل معلمة على نحو منفصل وإن طريقة المربعات الصغرى تتحطم , من جهة أخرى اذا لم ترتبط المتغيرات التوضيحية ببعضها ، بمعنى اذا كان معامل الإرتباط بين تلك المتغيرات يساوي صفراً فإن المتغيرات تصبح متغيرات متعامدة (Orthogonal) وليس هناك مشكلات تتعلق بتقدير المعلمات بقدر تعلق الأمر بمشكلة الإرتباط الخطي المتعدد ، في الواقع الفعلي في حالة المتغير المتعامد (Orthogonal Variable) (X'_s) ليس هناك حاجة لإجراء تحليل الإنحدار المتعدد ، حيث أن كل معلمة يمكن أن تقدر بوساطة تحليل الإنحدار البسيط (Simple Regression) مع المتغير (Y) بصيغة :

$$Y = F(X_i)$$

ولكن في الممارسة العملية أو في الواقع العملي ليس هناك أي من الحالتين المتطرفتين (وهو إرتباط خطي متعدد وعدم وجود إرتباط خطي متعدد) في الغالب الأعم ، في معظم الحالات هناك درجة من الإرتباط (Intercorrelation) بين المتغيرات التوضيحية يعود الى الإعتماد المتبادل (Interdependence) بين المتغيرات أو بين حجوم إقتصادية عبر الزمن ، في

هذه الحالة فإن معامل الإرتباط البسيط لكل زوج من المتغيــرات التوضــيحية سوف يأخذ قيمة بين الصفر والواحد وإن مشكلات الإرتباط المتعدد ربما تمنــع الدقة والإستقرار في المعلمــات المقــدرة (Parameter Estimates) ولكــن التأثيرات التامة للإرتباط المتعدد ولم يتم التأكد نظرياً منها بعد .

إن الإرتباط الخطي المتعدد ليس حالة أما أن تكون موجــودة أو غيــر موجودة في الدوال الإقتصادية ولكنها ظاهرة متأصلة (inherent) أو ملازمــة في معظم العلاقات ، وذلك يعود الى طبيعة الحجوم الإقتصادية . لــيس هنــاك دليل حاسم يتعلق بدرجة الإرتباط الخطي المتعدد التي إن وجدت فإنهــا ســوف تؤثر على نحو خطير في قيم المعلمات المقدرة . فمن البــديهي عنــدما يتغيــر متغيران توضيحيان آنياً وبالطريقة نفسها يصبح من الصعب جداً قيــاس تــأثير أحدهما على المتغير المعتمد على نحو منفصل ، على سبيل المثال إفتــرض إن الإنفاق الإستهلاكي لفرد من الأفراد يعتمد على دخله والأرصدة السائلة لديــه ، فإذا كان دخله وأرصدته السائلة خلال فترة من الزمن يتغيران بالنسبة نفســها ، فإن تأثير أحدهما على الإستهلاك ربما يعزى خطأ الى المتغير الآخر ، لذلك فإن تأثيرات هذين المتغيرين في الإستهلاك لا يمكن بحثها بسبب وجــود الإرتبــاط العالي المتبادل (Intereorrelation) بينهما .

9-2 أسباب ظهور أو وجود مشكلة الإرتباط الخطي المتعدد

إن مشكلة الإرتباط الخطي المتعدد ربما تظهر لأسباب مختلفة :

أولاً : ثمة ميل للمتغيرات الإقتصادية للحركة بعضها مع بعــض عبــر الزمن ، إن الحجوم الإقتصادية تتأثر بالعناصر نفسها ، وبالنتيجــة فــإن هــذه العناصر التي تقرر الحجوم الإقتصادية عندما تبــدأ فعلهــا أو تأثيرهــا فــإن المتغيرات أو الحجوم الإقتصادية تبين أو تظهر بنموذج سلوك واسع متشابه عبر الزمن . مثلاً في فترات الإزدهار أو النمو الإقتصادي فإن الحجوم الإقتصــادية

الأساسية تنمو على الرغم من أن بقية الحجوم تتباطيء خلفهـــا ، وهكـــذا فـــإن الدخل والإستهلاك والإدخار والإستثمار والأسعار والإستخدام تميل الى الإرتفاع في فترات التوسع الإقتصادي وتميل الى الإنخفاض في فترات الركود ، إن النمو وعناصر الإتجاه الزمني في السلاسل الزمنيـــة (Time Series) هـــي أكثـــر الأسباب خطورة في وجود أو ظهور مشكلة الإرتباط الخطي المتعدد .

ثانياً : إستخدام القيم المتباطئة أو المتخلفة زمنيـــاً (Lagged Values) لبعض المتغيرات التوضيحية بوصفها عناصر مستقلة في العلاقة ، فالنماذج ذات التباطؤات أو التخلفات الموزعة (Distributed Lags) كانت قد أعطت نتـــائج مرضية أو جيدة في عدة حقول من الإقتصاد القياسي كمـــا أن إســـتخدام تلـــك النماذج يتوسع بسرعة .

مثلاً في دوال الإستهلاك أصبح إعتيادياً إدخال مستوى الـــدخل للفتـــرة السابقة أو الماضية جنباً الى جنب مع مستوى الدخل في الفترة الحالية مع بقيـــة المتغيرات التوضيحية . وبالطريقة نفسها في دوال الإستثمار فإن التبـــاطوء أو التخلفات الموزعة التي تتعلق بمستويات النشاط الإقتصادي في الفترة السابقة قد أدخلت بوصفها متغيرات توضيحية ، أنه من الطبيعي أن تكون القيم المتتابعـــة لمتغير معين ترتبط بعضها ببعض ، مثلاً الدخل في الفترة الحالية يقرر جزئيـــاً بقيمته في الفترة السابقة ، وهكذا فإن الإرتباط الخطي المتعدد بوصفه مشكلة في تحليل الإنحدار هي غالباً موجودة بالتأكيد في نماذج التباطؤ الموزع ، فإذا أخذنا بنظر الإعتبار كل الأسباب المذكورة آنفاً والتي تسبب وجود الإرتباط الخطـــي المتعدد فإنه يبدو واضحاً أن درجة معينة من الإرتباط الخطي المتعدد تظهر في معظم العلاقات الإقتصادية .

يجب أن نلاحظ أنه على الرغم من أن الإرتباط الخطي المتعـــدد عـــادة يكون موجوداً في بيانات السلاسل الزمنية ، ولكن في بعـــض الأحيـــان يكـــون موجوداً في بيانات المقطع العرضي (Cross Section data) أيضاً ، مثلاً في

عينة من بيانات المقطع العرضي للشركات الصناعية فإن العمل ورأس المـــال بوصفهما عناصر إنتاج غالباً ما يرتبطان ببعضهما إرتباطـــاً قويـــاً وذلـــك لأن الشركات الكبيرة تميل الى أن يكون لها كميات كبيرة من العنصـــرين ، بينمـــا الشركات الصغيرة عادة يكون لها كميات أصغر من العمل ورأس المال ، ولكن وجود الإرتباط الخطي المتعدد يميل الى أن يكون مشكلة عامة وأكثر خطـــورة في السلاسل الزمنية .

9-3 ما يترتب على وجود الإرتباط الخطي المتعدد

اذا كان الإرتباط الخطي بين المتغيرات التوضيحية إرتباطاً عاماً أي أن $(r_{xixj} = 1)$ عندئذ :

أ- فإن المعلمات المقدرة تكون غير نهائية .

البرهان :إفترض إن العلاقة الآتية مطلوب تقديرها

$$Y = b_0 + b_1X_1 + b_2X_2 + u$$

ونفترض أن $(X_2)(X_1)$ مرتبطان ببعضهما بعلاقة تامة أي أن :

$$X_2 = kX_1$$

عندما يكون k أي رقم ثابت إعتباطي .

إن قانوني تقدير معلمات هذه العلاقة $(\hat{b}_2 ، \hat{b}_1)$ هما :

$$\hat{b}_1 = \frac{\Sigma(X_1-\bar{X}_1)(Y-\bar{Y})\Sigma(X_2-\bar{X}_2)^2 - \Sigma(X_2-\bar{X}_2)(Y-\bar{Y})\Sigma(X_1-\bar{X}_1)(X_2-\bar{X}_2)}{\Sigma(X_1-\bar{X}_1)^2\Sigma(X_2-\bar{X}_2)^2 - (\Sigma(X_1-\bar{X}_1)(X_2-\bar{X}_2))^2}$$

$$\hat{b}_2 = \frac{\Sigma(X_2-\bar{X}_2)(Y-\bar{Y})\Sigma(X_1-\bar{X}_1)^2 - \Sigma(X_1-\bar{X}_1)(Y-\bar{Y})\Sigma(X_1-\bar{X}_1)(X_2-\bar{X}_2)}{\Sigma(X_1-\bar{X}_1)^2\Sigma(X_2-\bar{X}_2)^2 - (\Sigma(X_1-\bar{X}_1)(X_2-\bar{X}_2))^2}$$

نعوض عن (X_2) بقيمة (k_{X1}) فنحصل على :

$$\hat{b}_1 = \frac{K^2\Sigma\left(X_1 - \overline{X}_1\right)\left(Y - \overline{Y}\right)\Sigma\left(X_1 - \overline{X}_1\right)^2 - K^2\Sigma\left(X_1 - \overline{X}_1\right)\left(Y - \overline{Y}\right)\Sigma\left(X_1 - \overline{X}_1\right)^2}{K^2\left(\sum\left(X_1 - \overline{X}_1\right)^2\right)^2 - K^2\left(\sum\left(X_1 - \overline{X}_1\right)^2\right)^2} = \frac{0}{0}$$

$$\hat{b}_2 = \frac{K\Sigma\left(X_1 - \overline{X}_1\right)\left(Y - \overline{Y}\right)\Sigma\left(X_1 - \overline{X}_1\right)^2 - K\Sigma\left(X_1 - \overline{X}_1\right)\left(Y - \overline{Y}\right)\Sigma\left(X_1 - \overline{X}_1\right)^2}{K^2\left(\sum\left(X_1 - \overline{X}_1\right)^2\right)^2 - K^2\left(\sum\left(X_1 - \overline{X}_1\right)^2\right)^2} = \frac{0}{0}$$

ولذلك فإن معلمات العلاقة $(\hat{b}_2 ، \hat{b}_1)$ هي معلمات غير نهائيـــة ولـــيس هناك أية طريقة لإيجاد قيم منفصلة لكل معلمة من المعلمات .

ب- إن الخطأ المعياري للمعلمات يصبح كبيراً على نحو لانهائي .

اذا كان $(r_{xixj} = 1)$ ، فإن الخطأ المعياري للمعلمات المقدرة يكون كبيراً على نحو لانهائي :

البرهان : في نموذج المتغيران :

$$Y = b_0 + b_1X_1 + b_2X_2 + u$$

فإذا كان (X_1) مرتبط على نحو تام بالمتغير التوضيحي الثاني (X_2) أي أن $(X_2=k_{X1})$ فإن تباينات (Variances) المعلمات المقدرة سوف تكون :

$$Var(\hat{b}_1) = \hat{\delta}_u^{\,2}\frac{\sum\left(X_2 - \overline{X}_2\right)^2}{\Sigma\left(X_1 - \overline{X}_1\right)^2\Sigma\left(X_2 - \overline{X}_2\right)^2 - \left(\Sigma\left(X_1 - \overline{X}_1\right)\left(X_2 - \overline{X}_2\right)\right)^2}$$

$$Var(\hat{b}_2) = \delta_u^{\,2}\frac{\sum\left(X_1 - \overline{X}_1\right)^2}{\Sigma\left(X_1 - \overline{X}_1\right)^2\Sigma\left(X_2 - \overline{X}_2\right)^2 - \left(\Sigma\left(X_1 - \overline{X}_1\right)\left(X_2 - \overline{X}_2\right)\right)^2}$$

نعوض عـــن (X_2) بقيمة (k_{X1}) فنحصل على :

$$Var(\hat{b}_1) = \hat{\delta}_u^{\,2}\frac{k^2\sum\left(X_2 - \overline{X}_2\right)^2}{k^2\Sigma\left(X_1 - \overline{X}_1\right)^2\Sigma\left(X_1 - \overline{X}_1\right)^2 - k^2\left(\Sigma\left(X_1 - \overline{X}_1\right)\right)^2}$$

$$= \frac{\hat{\delta}_u^{\,2}\Sigma\left(X_1 - \overline{X}_1\right)^2}{0} = \infty$$

وهكذا فإن تباينات المعلمات المقدرة تصبح لانهائية مالم يكن $(\delta_u^{\,2} = 0)$

ولكن على كل حال ليس هناك سبب مسبق ليميل (δ_u^2) الى أن يكــون صــفراً عندما يزداد الإرتباط المتعدد بين المتغيرات التوضيحية ، لتوضيح المشكلة لنأخذ المثال الآتي لعلاقة إقتصادية تتضمن ثلاثة متغيرات توضيحية ، لنفتــرض إن الإستهلاك الحقيقي لبلد معين هو :

$$Y = b_0 + b_1X_1 + b_2X_2 + b_3X_3 + u$$

عندما يكون :

(Y) = الإستهلاك الكلي

(X_1) = دخل المناطق الريفية

(X_2) = دخل المناطق الحضرية

(X_3) = الضريبة على الدخل

على أسس مسبقة نتوقع أن (b_1) أصغر من (b_2) طالما أن الميل الحدي للإستهلاك في المناطق الحضرية على نحو عــام أكبــر مــن الميــل الحــدي للإستهلاك (MPC) في المناطق الريفية .

نحن نرغب بالحصول على تقديرات للميــول الفرديــة ، كــذلك بقيــة المعلمات (parameters) . لنفترض أنه خلال الفتــرة التــي تغطيهــا العينــة (X_2,X_1) متساويان (بمعنى ان الدخل يوزع بالتساوي بين المنــاطق الريفيــة والمناطق الحضرية) وهذا يعني ($X_2=X_1$) . في ظل هذه الظروف سوف يكون من غير الممكن ان نحصل على تقديرات مستقلة للمعلمات (b_2,b_1) وذلــك لأن (X_2,X_1) يتغيران مع بعض ، وربما نعوض عن(X_2) بـــ(X_1) فنحصل على :

$$Y = b_0 + b_1X_1 + b_2X_1 + b_3X_3 + u$$

$$Y = b_0 + (b_1 + b_2)\ X_1 + b_3X_3 + u$$

وهكذا فإن إسقاط واحد من المتغيرين المتساويين يعطينا تقديراً لمجموع المعلمتين وليس قيماً فردية لكل من (b_2,b_1) وهذا يعني أن مجمــوع (b_1+b_2) سوف يكون مشخصاً ، ولكن (b_2,b_1) غيــر مشخصــة علــى نحــو مســتقل

(هناك علاقة قريبة بين الإرتباط الخطي المتعدد والتشـــخيص الإحصـــائي) اذا كانت قيم (X) غير مرتبطة ببعضها إرتباطاً تاماً ولكن هناك درجة من الإرتباط أي أن $(0 < r_{X1Xj} < 1)$ فإن تأثيرات الإرتباط الخطـــي غيـــر مؤكـــدة ، إن الدلائل من الدراسات الإقتصادية القياسية النظرية مع بيانات مســـيطر عليهـــا ، وكذلك من البحث التطبيقي هي دلائل تخضع للجدل وهي غير حاسمة أو غيـــر مقنعة .

في بعض الدراسات فـــإن القـــيم المعلمـــات تصـــبح غيـــر مســـتقرة (Unstable) كلما تم إدخال متغيرات مرتبطة إضافية مرتبطة بعضها الى الدالة أو كلما ازداد حجم العينة ، وفي دراسات أخرى فإن قيم تقديرات المعلمـــات لا تتأثر كثيراً أو تأثيراً مهماً ، والشيء نفسه بالنسبة للأخطاء المعيارية للتقديرات :

ففي بعض الدراسات وجد أن بعض الأخطاء المعيارية للمعلمات المقدرة تزداد أو ترتفع على نحو كبير عندما يكون في الدالة متغيرات مرتبطة بعضـــها البعض (Collinear) ، بينما في أمثلة أخرى وجد أن الأخطاء المعياريـــة لـــم يحصل أن تأثرت بالإرتباط الخطي المتعدد ، وهنا لابد لنا أن نؤكد على نقطتين أساسيتين :

<u>أولاً</u> : إن المعلمات المقدرة غير متحيزة إحصائياً $((E(\hat{b}_i) = b_i))$ حتى ولو كان الإرتباط الخطي المتعدد (Multicollinearity) قويـــاً ، إن الصـــفة الإحصائية في عدم التحيز (Unbiasedness) في تقـــدير المعلمـــات بطريقـــة المربعات الصغرى الإعتيادية (OLS) لاتتطلب أن قـــيم (X) غيـــر مرتبطـــة ببعضها .

من ناحية أخرى فإن العينات اتي تكون فيها علاقة إرتباط خطي متعدد بين قيم (X) ربما تجعل المعلمات المقدرة غير دقيقة وغير مستقرة ، ولســـوء الحظ ليس هناك قواعد راسخة لتقييم خطورة مثل تلك الأخطاء ، ولكـــن عـــدم إستقرار القيم المقدرة للمعلمات ربما يكون خطيراً جداً الى الحد الذي يؤدي فيه

الى تغيير إشارة المعلمات المقدرة كلما ازدادت درجة الإرتباط الخطي ، هنـــاك بعض الدلائل تشير الى أن زيادة الإرتباط الخطي المتعدد ينتج عنها تغيرات في قيم المعلمات إعتماداً على أهمية كل متغير توضيحي ، والأهمية تقاس عادة من خلال معامل الإرتباط البسيط بين (Y) وكل من قيم (X) (r_{yxi}) .

إن مثل هذه الدلائل كانت قد قادت عدداً كبيراً مـــن المختصـــين فـــي الإقتصاد القياسي الى القول إن تأثيرات الإرتباط الخطي المتعدد على المعلمات المقدرة تعتمد على درجة قساوة الإعتماد المتبادل (interdependence) بـــين المتغيرات وعلى أهمية تلك المتغيرات التي تدخل في علاقات خطية مع بعضها البعض ، في أي دراسة معينة هناك متغيرات توضـــيحية اكثـــر أهميـــة مـــن المتغيرات الأخرى ، مثلاً في دالة الطلب فإن سعر السلعة ودخـــل المســـتهلك وسعر البدائل القريبة من السلعة هي متغيرات أكثر أهمية من أســـعار البـــدائل البعيدة للسلعة على أسس مسبقة بالنسبة للسلعة قيد الدراســـة ، اذا كانـــت تلـــك المتغيرات التوضيحية حاسمة ستراتيجياً وحدث وأنها مرتبطة ببعضها بقوة فإن خطورة مشكلة الإرتباط الخطي المتعدد تكون أكبر مما في حالـــة كـــون تلـــك المتغيرات المرتبطة ببعضها تشكل عناصر ثانوية وذلك لأن العناصر الثانويـــة يمكن إسقاطها من التحليل دون التأثير على النتائج على نحو خطير .

ثانياً : يبدو أن كتّاب عدة يقبلون أنه عندما يكون الإرتبـــاط الخطـــي المتعدد موجوداً في دالة فإن الأخطاء المعيارية للمعلمات المقدرة سوف تكـــون على نحو عام كبيرة ، ولكن هذا ليس صحيحاً دائماً وذلك لأن كلاً من البســـط والمقام لقانون حساب التباينات سوف تكون متأثرة عادة بالصيغ التي تتضـــمن المجموع من منتجات (X_s) لذلك فإن الحجم النهائي لتباين (b_i) ربما لا يكـــون كبيراً .

يبدو أن لورنس كلاين (L.R.Klein) يقبل بأن الإرتباط الخطي المتعدد ليس بالضرورة مشكلة مالم يكن ذلك الإرتباط الخطي المتعدد أكبر نسبياً من درجة الإرتباط المتعدد الكلي (Overall) بين كل المتغيرات آنياً .

يعتقد كلاين إن الإرتباط الخطي يكون مؤذياً اذا كان :

$$r^2_{xixj} \geq R^2_{y.x1.x2} \ldots\ldots\ldots x_n$$

عندما يكون :

(r^2_{xixj}) = معامل الإرتباط البسيط بين أي متغيرين توضيحيين (x_j,x_i) .

(R^2) = الإرتباط الكلي للعلاقة التي يمثلها خط الإنحدار .

إن إسلوب كلاين هوجم من قبل كل من فرار (Farrar) وكلاوبر (Glauber) في بحث نشر في مجلة أمريكية "مشكلة الإرتباط الخطي المتعدد في تحليل الإنحدار" " Multicdlinearity in Regression Analysis" نشر في مجلة عام 1967 " Review of Economics and Statistics " .

ومن ناحية أخرى يعتقد ثايل (Theil) إن النموذج مع أكثر من متغيرين من (X) حتى لو كانت الإرتباطات بين المتغيرات التوضيحية صغيرة ربما تقود الى عدم أهمية المتغيرات إحصائياً وذلك يعود الى زيادة الأخطاء المعيارية ، ولكن لابد لنا ان نقول أن فريش (Frisch) أوضح أن الأخطاء المعيارية هي ليست كبيرة دائماً عند وجود الإرتباط الخطي المتعدد ، وهكذا ربما نحصل على تقديرات للمعلمات غير دقيقة بسبب الإرتباط الخطي المتعدد على الرغم من أن الأخطاء المعيارية لتلك التقديرات الخاطئة ربما لا تبينها .

ولتلخيص المناقشات ربما نقول أنه على الرغم من وجود إستثناءات على نحو عام فإن زيادة الأخطاء المعيارية تظهر عندما ندخل متغيرات ترتبط ببعضها بوصفها متغيرات توضيحية في الدالة .

وهكذا مع وجود الإرتباط الخطي المتعدد في الدالة نواجه خطر الخطــأ في التحديد وذلك بسبب أنه ربما نرفض المتغير الذي يظهر فيه الخطأ المعياري عالياً على الرغم من أن ذلك المتغير هو مقرراً مهماً للإنحرافات في المتغيــر المعتمد .

9-4 إختبارات للكشف عن وجود الإرتباط الخطي المتعدد

9-4-1 طريقة معتمدة على تحليل الحشد (طريقة فريش Frisch)

إن خطورة تأثير الإرتباط الخطي المتعدد تبدو معتمــدة علــى درجــة الإرتباطالتبادل بين المتغيرات (Intercorrelation)(r_{XiXj}) التوضيحية ، وكذلك معامل الإرتباط الكلي ($RY.X_1X_2...X_n$) معامل الإرتباط للإنحدار ، وهكــذا يمكن أن نقول إن الأخطاء المعيارية (standard Errors) ومعاملات الإرتباط الجزئي (partial Correlation Coefficient)(r_{XiXj}) ومعامل الإرتباط الكلي (R^2) يمكن أن تستعمل لإختبار وجود الإرتباط الخطي المتعدد ، ولكن أي مــن هذه المعايير بنفسه لا يعد مئشراً كافياً أو مرضياً للإرتباط الخطي المتعدد وذلك يعود الى الأسباب التالية :

أ- إن الأخطاء المعيارية الكبير لاتظهر مع الإرتباط الخطــي المتعــدد (من الدلائل ما يخص دالة الإنتاج بصيغة كوب دوكــلاس (Cobb-Douglas) والأكثر من هذا إن الأخطاء المعيارية الكبيرة ربما تظهر لأسباب مختلفة وليس فقط بسبب وجود علاقة خطية بين المتغيرات التوضيحية .

ب- إن الإرتباطات بين المتغيرات التوضيحية ليس بحاجة الى ان تكون عالية حتى تتأثر قيم المعلمات (b_i) والأخطاء المعيارية لها تأثيراً سلبياً أو رديئاً، وهذا يعني أن (r_{XiXj}) ليس معياراً كافياً بحد ذاته .

جـــ- إن معامل الإرتباط الكلي (R^2)(Overall) ربما يكون عالياً نسبة الى معاملات الإرتباط بين المتغيرات التوضيحية (r_{XiXj}) ولكن ربما تبقى النتائج غير دقيقة على نحو كبير وليس ذات معنوية إحصائية ، (مع إشارات خاطئــة

وأخطاء معيارية كبيرة) ، ولكن على كل حال فإن تركيبة أو توليفة (a combination) من هذه المعايير جميعها ربما تساعد في الكشف عن الإرتباط الخطي المتعدد ، ومن أجل الحصول على القدر الممكن من المعرفة حول خطورة مشكلة الإرتباط الخطي المتعدد نحاول أن نتبنى طريقة هي في جوهرها صيغة منقحة لطريقة فريش في تحليل الحشد أو طريقة تحليل خارطة الحزمة (Bunch-Map Analysis) ، والأسلوب هو وضع صيغة إنحدار بين المتغير المعتمد على كل متغير توضيحي على نحو مستقل ، وهكذا نحصل على إنحدارات أولية أو بسيطة ونقوم بفحص أو إختبار نتائجها على أساس معايير مسبقة ومعايير إحصائية ، ونقوم بإختيار الإنحدار البسيط الذي يبدو أنه يعطينا النتائج الأكثر إحتمالاً أو حصولاً أو إمكانية الحصول على أسس مسبقة وأسس إحصائية ، ومن ثم ندخل تدريجياً متغيرات إضافية ونختبر تأثيرها على المعاملات المنفردة وعلى الأخطاء المعيارية وقيمة (R^2) .

إن المتغير الجديد المضاف يصنف أنه مفيد أو إنه متغير غير ضروري أو زائد (Superflous) أو متغير ضار أو مؤذي أو غير مرغوب فيه أو معيق أو كما يأتي :

1- اذا أدى المتغير الجديد الى التحسن في (R^2) دون أن يجعل المعاملات أو المعلمات غير مقبولة أو مخطوءة على أساس إعتبارات مسبقة ، فإن ذلك المتغير يعد أو يصنف أنه مفيد ويحتفظ به بوصفه متغيراً توضيحياً .

2- اذا لم يؤدي المتغير الجديد الى التحسن في (R^2) ولم يؤثر على قيم المعلمات المفردة الى مدى كبير او مهم ، فإن المتغير يعد غير ضروري أو زائد ، ويجب أن يرفض بمعنى عدم ضمه الى مجموعة المتغيرات التوضيحية .

3- اذا أثر المتغير الجديد على إشارات وحجوم المعلمات الى مدى كبير فإن ذلك المتغير يعد ضاراً أو مؤذياً أو غير مرغوب فيه ، فإذا تأثرت المعلمات بطريقة بحيث تصبح غير مقبولة على أسس أو إعتبارات نظرية مسبقة ، عندئذ

ربما نقول بأن هذا يعد تحذيراً أن مشكلة الإرتباط الخطي المتعدد في هذه الحالة الخطيرة ، إن المتغير الجديد مهم لكن بسبب الإرتباطات المتبادلة مع المتغيرات المستقلة الأخرى فإن تأثيره لا يمكن أن يقــوم إحصـــائياً بطريقـــة المربعـــات الصغرى الإعتيادية (Ordinary Least Squares) .

إن هذا لا يعني أننا نرفض أو يجب أن نـــرفض المتغيـــر المـــؤذي أو الضار ، فإذا فعلنا ذلك فإننا سوف نهمل معلومات قيمة بالنسبة الى محاولاتنا في الوصول الى أفضل تحديد للعلاقة ، فإذا حذفنا المتغير المؤذي أو الضار تمامـــاً في محاولة لتجنب أثره المؤذي على بقية المعلمات فإننا يجـــب أن نضـــع فـــي تفكيرنا أن هذه العملية ببساطة تترك تأثير ذلك المتغير ليتم إمتصاصه من قبـــل بقية المعلمات الأخرى (بحيث تكون قيمتها مختلطـــة Mixed) ، وإمتصاصـــه أيضاً قبل المتغير العشوائي الذي ربما يصبح مرتبطاً مع بقية المتغيرات التـــي تركت في الدالة ، وبالنتيجة مخالفة الفرض السادس لأنه في هذه الحالة فإن :

$$E\ (UiXj \neq 0)$$

مثال : الجدول (9-1) يحتوي على بيانات السلاســـل الزمنيـــة للفتـــرة 1951-1968 حول الإنفاق على الملابس والدخل تحت التصرف والأرصـــدة السائلة والرقم القياسي لأسعار الملابس والرقم القياسي العـــام للأســـعار لأحـــد البلدان ، والمطلوب في هذا المثال تقدير دالة الطلب على الملابس .

جدول (9-1)

السنة	الإنفاق على الملابس	الدخل تحت التصرف	الأرصدة السائلة	الرقم القياسي لأسعار الملابس 1963=100	الرقم القياسي العام للأسعار 1963=100
1959	8.4	82.9	17.1	92	94
1960	9.6	88.0	21.3	93	96
1961	10.4	99.9	25.1	96	97
1962	11.4	105.3	29.0	94	97
1963	12.2	117.7	34.0	100	100
1964	14.2	131.0	40.0	101	101
1965	15.8	148.2	44.0	105	104

1966	17.9	161.8	49.0	112	109
1967	19.3	174.2	51.0	112	111
1968	20.8	184.7	53.0	112	111

على أسس نظرية مسبقة نقول إن الإنفاق على الملابس يتأثر بكل العناصر الموجودة بياناتها في الجدول (1) لذلك فإن دالة الطلب على الملابس يجب أن تأخذ الشكل الآتي :

$$C = b_0 + b_1Y + b_2L + b_3P_c + b_4P_o + u$$

عندما يكون :

(C) = الإنفاق على الملابس .

(L) = الأرصدة السائلة .

(P_c) = سعر الملابس .

(Y) = الدخل تحت التصرف

(P_o) = سعر السلع الأخرى .

وبتطبيق طريقة المربعات الصغرى على هذه الدالة نحصل على التقديرات الآتية :

$$\underset{(7.5)}{C = -13.53} + \underset{(0.03)}{0.09Y} + \underset{(0.05)}{0.015L} + \underset{(0.09)}{0.199P_c} + \underset{(0.15)}{0.34P_o}$$

$$R^2 = 0.998 \qquad \sum \hat{y}^2 = 28.15 \qquad \sum \ell_i^2 = 0.33 \qquad d = 3.4$$

وبتطبيق تحليل التباين لإختبار جودة توفيق الإنحدار :

$$F^* = \frac{\dfrac{\sum \hat{y}^2}{(k-1)}}{\dfrac{\sum \ell_i^2}{(n-k)}} = \frac{\dfrac{28.15}{4}}{\dfrac{0.33}{5}} = 15.6$$

طالما أن القيمة النظرية لــ(F) عند مستوى معنوية (0.05) مع درجات حرية(υ_1=k-1 =4) (υ_2=n- k=5) هي (5.19) نرفض فرضية العدم ونقبل الفرضية البديلة التي تنص على أن هناك علاقة ذات معنوية إحصائية عالية بين

الإنفاق على الملابس والمتغيرات التوضيحية ، ولكن كل المتغيرات التوضيحية مرتبطة ببعضها خطياً على نحو خطير وكما يبدو ذلك من معاملات الإرتبـــاط البسيط :

$r_{yL}= 0.993$ $r_{ypc} = 0.980$ $r_{ypo} = 0.987$

$r_{LPc} = 0.964$ $r_{LPo} = 0.973$ $r_{PcPo} = 0.991$

للكشف عن تأثيرات الإرتباط الخطـــي المتعـــدد نحســـب الإنحـــدارات البسيطة :

1- $\hat{C} = \hat{a}_0 = \hat{a}_1 y$
$= -1.24 + 0.118Y$ $r^2=0.995$ $d=2.6$
(0.38) (0.002)

2- $\hat{C} = \hat{b}_0 = \hat{b}_1 Pc$
$= -38.51 + 0.516Pc$ $r^2=0.951$ $d=2.4$
(4.20) (0.04)

3- $\hat{C} = \hat{c}_0 = \hat{c}_1 L$
$= 2.11 + 0.327L$ $r^2=0.967$ $d=0.4$
(0.81) (0.02)

4- $\hat{C} = \hat{d}_0 = \hat{d}_1 Po$
$= -53.65 + 0.663Po$ $r^2=0.977$ $d=2.1$
(3.65) (0.03)

نختار الإنحدار الأول [C=f(Y)] بوصفه خطوة أولى في التحليل طالما أن الدخل (Y) يبدو على أسس نظرية مسبقة هو المتغير التوضـــيحي الأكثـــر أهمية خلال المدة الزمنية قيد الدراسة ، عندئذ ندخل بقية المتغيرات التوضيحية تدريجياً في الدالة والنتائج تبيد في الجدول (9−2) :

جدول (9-2)

الدالة	$\hat{b}_0$ ثابت	$\hat{b}_1$ (Y)	$\hat{b}_2$ (P_c)	$\hat{b}_3$ (L)	$\hat{b}_4$ (P_0)	R^2	d
C=f(Y)	-1.24	0.118				0.995	2.6
	(0.37)	(0.002)					
$C=f(Y,P_c)$	1.40	0.126	-0.036			0.996	2.5
	(4.92)	(0.01)	(0.07)				
$C=f(Y,P_c,L)$	0.94	0.138	-0.034	-0.37		0.996	3.1
	(5.17)	(0.02)	(0.06)				
$C=f(Y,P_c,P_0)$	-12.76	0.104	-0.188		0.319	0.997	3.5
	(6.52)	(0.01)	(0.07)		(0.12)		
$C=f(Y,P_c,L,P_0)$	-13.53	0.097	-0.99	0.015	0.34	0.998	3.4
	(7.5)	(0.03)	(0.09)	(0.05)	(0.15)		

* ملحوظة : الأرقام بين الأقواس هي الأخطاء المعيارية للمعلمات المقدرة .

إن التغيرات في الدخل تبدو هي الأكثر أهمية في توضيح الإنحرافات في الإنفاق على الملابس ، إن إدخال المتغير (P_c) حسن معامل التحديد (R^2) على نحو بسيط ، إن إشارات ($\hat{b}$) صحيحة ولكن الأخطاء المعيارية تبين أن ($\hat{b}_2$) ليست ذات معنوية إحصائية .

إن الإرتباط المتبادل العالي بين (P_cوY) لا يؤثر على إستقرار أو معنوية ($\hat{b}_1$) .

إن إدخال الأرصدة السائلة لا يعطي تقديرات جيدة لـ(b_2) أو (b_3) ، واضح من التقدير أن الإرتباط المتبادل العالي بين (P_c و L) يجعل من المستحيل الحصول على تقديرات ذات معنى منفصلة لكل من (b_2) و (b_3) ولكن تقدير المعلمة ($\hat{b}_1$) غير متأثر على الرغم من وجود الإرتباط المتبادل الخطير بين (Y , P_c , L) ، وهكذا فإن (L) ربما يعد متغير غير ضروري أو

زائد ، وبإسقاط المتغير (L) وإدخال المتغير (P_o) في الدالة نحصل على توفيق جيد ، أي معامل تحديد (R^2) أفضل أو يزداد قليلاً ، وإن كل المعلمات المقدرة أخذت الإشارة الصحيحة وهي ذات معنوية إحصائية ، ولكن على الرغم من الدرجة العالية من الإرتباط الخطي بين كل المتغيرات التوضيحية فإن الأخطاء المعيارية ليست كبيرة .

إن الإنحدار مع كل المتغيرات التوضيحية الأربعة يبين أن تأثير الإرتباط الخطي المتعدد غير خطير بالنسبة الى ($\hat{b}_1$ $\hat{b}_2$) أما معلمة المتغير (L)($\hat{b}_3$) فهي ليست ذات معنوية إحصائية ، وهكذا فإن المتغير (L) بوضوح هو متغير زائد أو غير ضروري ، اذاً فإن أفضل توفيق يمكن الحصول عليه من الدالة الآتية :

$$\mathbf{C = f(Y,P_c,P_0)}$$

9-4-2 إختبار فرار - كلاوبر (Farrar-Glauber) للإرتباط الخطي المتعدد (Multicollinearity)

نشر فرار وكلاوبر في عام 1967 بحثاً في مجلة إقتصادية* حاولا فيه عرض إختباراً إحصائياً للإرتباط الخطي المتعدد ، وهو في الحقيقة عبارة عن منظومة من ثلاثة إختبارات ، حيث أن فرار وكلاوبر إستخدما ثلاثة أنواع من الإحصاءة (Statistics) لإختبار الإرتباط المتعدد .

الإختبار المتعدد أو الإحصاءة الأولى هي مربع كاي (χ^2)(Chi-square) لإكتشاف وجود خطورة الإرتباط الخطي المتعدد في الدالة التي تتضمن متغيرات توضيحية عدة ، والإختبار الثاني هو إختبار (F) لإكتشاف موقع المتغيرات الداخلة في إرتباط خطي متعدد ، الإختبار الثالث هو إختبار (t) وهذا الإختبار لغرض إيجاد نمط الإرتباط الخطي المتعدد ، بمعنى

---

* **D.E.Farrar and R.R.Glauber,"Multicollinearity in Regression Analysis" Review of Economics and Statistics Vol.49,1967 PP (92-107)**

المساعدة في تقرير أي المتغيرات مسؤول عن ظهور الإرتباط الخطي المتعدد ، لقد عُد الإرتباط الخطي المتعدد في عينة ما من قبل فـــرار وكلاوبـــر بوصـــفه إبتعاداً للقيم المشاهدة لـــ(X_s) عـــن الأورثاكونـــالتي (Orthogonality) إذ إن أسلوبهما كان قد إنبثق من فكرة مفادها أنه اذا كان الإرتباط الخطي المتعدد تاماً، عندئذ فإن المعلمات المقدرة تصبح غير نهائية وأن علاقات الإرتباط المتبـــادل بين مختلـــف المتغيـــرات التوضـــيحية يمكـــن قياســـها بوســـاطة معـــاملات (Coefficient) الإرتباط المتعدد ومعاملات الإرتباط الجزئي ، إن إختبـــارات فرار ــ كلاوبر يمكن أن نصفها بالإطار العام الآتي :

أولاً ــ إختبار مربع كاي (Chi-square test)

لإختبار خطورة أو قساوة الإرتباط الخطي المتعدد في حالـــة تتضـــمن متغيرات توضيحية عدة ، إن الفرضية الموضوعة قيد الإختبار في هذه المرحلة هـــي إن جمـــع (X_s) فـــي العينـــة هـــي أورثاكونـــال (Orthogonal) (r_{xixj}=r and r_{xixj}=0) ، لهـــذا مـــن الملائـــم أن تقـــوم بعمليـــة معياريـــة (To Standardise) للمتغيرات لحجم العينة وللإنحراف المعياري ، إن العملية المعيارية تنفذ من خلال تقسيم كل المشاهدات لكل (X) والتي تعرض بصـــيغة الإنحرافات عن وسطها الحسابي على ($\sqrt{n}$) مضروباً في الإنحراف المعياري لـــ(X) ، وهذا يعني القيمة المعيارية لـــ(t^{th}) من المشـــاهدات لـــــ(j^{th}) مـــن المتغير:

$$\frac{\left(X_{jt}-\overline{X}_{j}\right)}{\sqrt{n}\left(S_{xj}\right)}$$

إن هذه العملية مساوية لتقسيم كل عنصر من عناصر محددة ، مجموع المربعات ومجموع منتجات (X)(معروضة بصيغة الإنحرافات) على الجـــذور

التربيعية لمجموع مربعات الإنحرافات للمتغيرات التي تظهر في العنصر ، مثلاً محددة (X) بصيغة الإنحرافات في نموذج الثلاث متغيرات هي :

$$\begin{vmatrix} \sum(X_1-\bar{X}_1)^2 & \sum(X_1-\bar{X}_1)(X_2-\bar{X}_2) & \sum(X_1-\bar{X}_1)(X_3-X_3) \\ \sum(X_1-\bar{X}_1)(X_2-\bar{X}_2) & \sum(X_2-\bar{X}_2)^2 & \sum(X_2-\bar{X}_2)(X_3-\bar{X}_3) \\ \sum(X_1-\bar{X}_1)(X_3-\bar{X}_3) & \sum(X_2-\bar{X}_2)(X_3-\bar{X}_3) & \sum(X_3-\bar{X}_3)^2 \end{vmatrix}$$

وللحصول على الشكل المعياري لهذه المحددة نقسم العنصر الأول فيها على $\left(\sqrt{\sum(X_1-\bar{X}_1)^2}\right)^2 = \sqrt{\sum(X_1-\bar{X}_1)^2}\sqrt{\sum(X_1-\bar{X}_1)^2}$ والعنصر الثاني على $\left(\sqrt{\sum(X_1-\bar{X}_1)^2}\sqrt{\sum(X_2-\bar{X}_2)^2}\right)$ وهكذا وعلى نحو عام فإن العنصر $\left(\sum(X_i-\bar{X}_i)(X_j-\bar{X}_j)\right)$ يقسم على $\left(\sqrt{\sum(X_i-\bar{X}_i)^2}\sqrt{\sum(X_j-\bar{X}_j)^2}\right)$ لإعطاء العنصر المرافق للمحددة المعيارية ، بالنسبة لنموذج الثلاث متغيرات فإن المحددة المعيارية هي كما يأتي :

$$\begin{vmatrix} \frac{\Sigma(x_1-\bar{x}_1)^2}{\sqrt{\Sigma(x_1-\bar{x}_1)^2}} & \frac{\Sigma(x_1-\bar{x}_1)(x_2-\bar{x}_2)}{\sqrt{\Sigma(x_1-\bar{x}_1)^2}\sqrt{\Sigma(x_2-\bar{x}_2)^2}} & \frac{\Sigma(x_1-\bar{x}_1)(x_3-\bar{x}_3)}{\sqrt{\Sigma(x_1-\bar{x}_1)^2}\Sigma(x_3-\bar{x}_3)^2} \\ \frac{\Sigma(x_1-\bar{x}_1)(x_2-\bar{x}_2)}{\sqrt{\Sigma(x_1-\bar{x}_1)^2}\sqrt{\Sigma(x_2-\bar{x}_2)^2}} & \frac{\Sigma(x_2-\bar{x}_2)^2}{\sqrt{\Sigma(x_2-\bar{x}_2)^2}} & \frac{\Sigma(x_2-\bar{x}_2)(x_3-\bar{x}_3)}{\sqrt{\Sigma(x_2-\bar{x}_2)^2}\sqrt{\Sigma(x_3-\bar{x}_3)^2}} \\ \frac{\Sigma(x_1-\bar{x}_1)(x_3-\bar{x}_3)}{\sqrt{\Sigma(x_1-\bar{x}_1)^2}\sqrt{\Sigma(x_3-\bar{x}_3)^2}} & \frac{\Sigma(x_2-\bar{x}_2)(x_3-\bar{x}_3)}{\sqrt{\Sigma(x_2-\bar{x}_2)^2}\sqrt{\Sigma(x_3-\bar{x}_3)^2}} & \frac{\Sigma(x_3-\bar{x}_3)^2}{\Sigma(x_3-\bar{x}_3)^2} \end{vmatrix}$$

إن المحددات المعيارية لمقامات تقديرات المربعات الصغرى يمكن إعادة كتابتها على نحو مختلف قليلاً على أن نتذكر أن عناصر الداياكنول (diagonal elements) تساوي واحد والعناصر الأخرى هي عبارة عن معاملات الإرتباط البسيط بين المتغيرات التوضيحية ، وهكذا فإن المحددة

المعيارية يمكن أن تدعى "محددة الإرتباط" وفي نموذج الثلاث متغيــرات فــإن المحددة المعيارية هي :

$$\begin{vmatrix} 1 & r_{X1X2} & r_{X1X3} \\ r_{X1X2} & 1 & r_{X2X3} \\ r_{X1X3} & r_{X2X3} & 1 \end{vmatrix}$$

ومن هذه الصيغ نستطيع أن نختبــر بســهولة الحــالتين المتطــرفتين الأورثوكونالتي (Orthogonality) والإرتباط الخطي المتعدد التام ، ففي حالة الإرتبــاط الخطــي المتعــدد التــام فــإن معــاملات الإرتبــاط البســيط ($r_{X1X2}, r_{X2X3}, \ldots ets$) تساوي واحد ولذلك فإن قيمة محددة الإرتباط المعياريــة تساوي صفراً ، في نموذج المتغيرين لدينا :

$$\begin{vmatrix} 1 & r_{X1X2} \\ r_{X1X2} & 1 \end{vmatrix} = \begin{vmatrix} 1 & 1 \\ 1 & 1 \end{vmatrix} = 0$$

وفي حالة الأورثوكونالتي (Orthogonality) لــ(X_S) فــإن معامــل الإرتباط البسيط لكل زوج من (X_S) يساوي صفراً ولذلك فــإن قيمــة محــددة الإرتباط المعيارية تساوي واحد ، في نموذج المتغيرين لدينا :

$$\begin{vmatrix} 1 & r_{X1X2} \\ r_{X1X2} & 1 \end{vmatrix} = \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 1$$

يتبع هذا اذا كانت قيمة المحددة المعيارية تقع بين الصفر والواحد فــإن هناك درجة من الإرتباط الخطي المتعدد .

إن التحليل المذكور آنفاً يقترح أن الإرتباط الخطي المتعدد ربمــا يعــد بوصفه إبتعاداً عن الأورثوكونالتي ، فكلما كان الإبتعاد عن الأورثوكونالتي قوياً

أي أن قيمة المحددة تقترب من الصفر فإن درجة الإرتباط الخطي المتعدد تكون أقوى والعكس بالعكس .

إبتداءً من هذه الحقيقة إقترح فرار وكلاوبر إختبار مربع كاي (χ^2) الآتي للكشف عن قوة الإرتباط الخطي المتعدد عبر منظومة المتغيرات التوضيحية برمتها أو كلها .

إن الفرضية الأساسية هي :

إن X كلها اورثاكونال (Orthogonal) :H_0

ضد الفرضية البديلة :

إن Xs ليس اورثاكونال (Orthogonal) :H_1

وجد فرار وكلاوبر إن الكمية :

$$*\chi^2 = -\left[n-1-\frac{1}{6}(2k+5)\right]\cdot \log_e [\text{قيمة المحددة المعيارية}]$$

عندما يكون :

$\chi^{2}*$ = القيمة المشاهدة لمربع χ^2 (حسب العينة)

n = حجم العينة الإحصائية

k = عدد المتغيرات التوضيحية

لها توزيع مربع (χ^2) مع درجات حرية تساوي $\left(\upsilon = \frac{1}{2}k(k-1)\right)$ من بيانات العينة الإحصائية نحصل على القيمة التجريبية لمربع كاي (χ^2*) والتي نقارنها مع القيمة النظرية لمربع كاي بمستوى معنوية مختار (يمكن الحصول على القيمة النظرية من جدول χ^2) مع درجات حرية $\left(\upsilon = \frac{1}{2}k(k-1)\right)$ ، يجب أن يكون واضحاً أن القيمة النظرية لمربع كاي (χ^2) هي القيمة التي تعرف المنطقة الحرجة (Critical Region) من الإختبار عند المستوى المختار من

المعنوية مع درجات الحرية الملائمة ، اذا كانت قيمة مربع كاي (χ^2*) المشاهدة أكبر من القيمة النظرية لمربع كاي (χ^2) مع درجات حرية $\left[\frac{1}{2}k(k-1)\right]$ نرفض إفتراض الأورثوكونالتي (Orthogonality) بمعنى أننا نقبل بأن هناك إرتباط خطي متعدد في الدالة وكلما كانت قيمة مربع كاي المشاهدة (χ^2*) عالية كلما كانت مشكلة الإرتباط الخطي المتعدد أكثر قساوة وخطورة .

أما اذا كانت القيمة المشاهدة لمربع كاي (χ^2*) أصغر من القيمة النظرية لها ($\chi^2* < \chi^2*$) فإننا نقبل إفتراض الأورثوكونالتي أي إننا نقبل بأنه ليس هناك مشكلة إرتباط خطي متعدد ذات معنوية في الدالة .

<u>ثانياً – إختبار (F) لتحديد موقع الإرتباط الخطي .</u>

من أجل إيجاد موقع العناصر التي تدخل في علاقة خطية متعددة قام فرار وكلاوبر بحساب معاملات الإرتباط المتعدد بين المتغيرات التوضيحية ($R^2_{Xi.X2X3\ldots Xk}, R^2_{X2,X1X3\ldots\ldots Xk}$) وعلى نحو عام($R^2_{Xi.X2X3\ldots Xk}$) وإختبار المعنوية الإحصائية بوصفها معاملات الإرتباط المتعدد مع إختبار (F) وكما يأتي :

لكل معامل إرتباط متعدد نحسب القيمة المشاهدة ل(F*) :

$$F* = \frac{(R^2_{X_i.X_1X_2\ldots\ldots X_k})/(k-1)}{(1-R^2_{X_i.X_1X_2\ldots\ldots\ldots X_k})/n-k)}$$

عندما يكون :

n = حجم العينة الإحصائية

k = عدد المتغيرات التوضيحية

الفرضية قيد الإختبار في هذه المرحلة هي :

$$H_0 : R^2_{Xi.X2X3\ldots Xk} = 0$$

والفرضية البديلة هي :

$$H_1 : R^2_{Xi.X2X3\ldots Xk} \neq 0$$

إن القيمة المشاهدة لـ(F*) تقارن مع القيمة النظرية لـ(F) (من جدول (F) مع درجات حرية $(\upsilon_1 = (k-1))$ $(\upsilon_2 = (n-k))$ عند مستوى معنوية مختار ، فإذا كانت قيمة (F*) أكبر (>) من (F) نقبل أن المتغير (X_i) يدخل في علاقة إرتباط خطي متعدد وهذا يعني أننا نرفض فرضية العدم ، اذا كانت قيمة (F*) أصغر (<) من (F) فإننا نقبل أن المتغير (X_i) غير داخل في إرتباط خطي متعدد .

ثالثاً – إختبار (t) لنمط الإرتباط الخطي المعتمد

يهدف إختبار (t) الى اكتشاف المتغيرات التي تسبب الإرتباط الخطي المتعدد ، ومن اجل إيجاد أي المتغيرات مسؤول عن الإرتباط الخطي المتعدد نحسب معاملات الإرتباط الجزئي بين المتغيرات التوضيحية وإختبار معنويتها الإحصائية مع إحصاءة (t) ، وللتذكير فقط نقول أن معامل الإرتباط الجزئي بين أي متغيرين (X_i) (X_j) يبين درجة الإرتباط بين المتغيرين مع بقاء كل المتغيرات الأخرى ثابتة ، وفي نموذج المتغيرين فإن معامل الإرتباط الجزئي هو معامل الإرتباط البسيط نفسه ، وفي نموذج الثلاث متغيرات فإن معاملات الإرتباط الجزئي تعطى بالصيغة الآتية :

$$r^2_{X1X2.X3} = \frac{(r_{12} - r_{13} - r_{23})^2}{(1 - r^2_{23})(1 - r^2_{13})}$$

$$r^2_{X1X3.X2} = \frac{(r_{13} - r_{12} - r_{23})^2}{(1 - r^2_{23})(1 - r^2_{12})}$$

$$r^2_{X2X3.X1} = \frac{(r_{23} - r_{12} - r_{13})^2}{(1 - r^2_{13})(1 - r^2_{12})}$$

أما بالنسبة للنماذج ذات المتغيرات الأكثر من ثلاثة متغيرات توضيحية فإن صيغة مماثلة يمكن أن توضع :

$$H_0 : r_{xixj....Xi.X2...Xk} = 0$$

تختبر هذه الفرضية مقابل الفرضية البديلة :

$$H_1 : r_{XiXj.X2X3...Xk} \neq 0$$

وعند إتمام تقدير معاملات الإرتباط الجزئي نحسب معنوياتها الإحصائية من خلال حساب إحصاءة (t^*) لكل واحد منها :

$$t^* = \frac{\left(rX_iX_j.X_1X_2.....X_k\right)\sqrt{n-k}}{\sqrt{1 = r_{X_iX_j.X_1X_2........X_k}}}$$

عندما يكون :

$r_{XiXj.X2X3...Xk}$ يشير الى معامل الإرتباط بين المتغير (X_i) والمتغير (X_j).

إن القيمة المشاهدة لــ(t^*) تقارن مع القيمة النظرية لــ(t) مع درجــات حرية ($\upsilon = (n-k)$) بمستوى مختار من المعنوية .

اذا كانت (t^*) أكبر من (t) الجدولية نقبل أن معامل الإرتباط الجزئــي بين المتغيرين (X_i) و (X_j) ذات معنوية إحصائية وهذا يعنــي ان المتغيــرين (X_i) و (X_j) هما مسؤولان عن الإرتباط الخطي المتعدد .

اذا كانت قيمة (t^*) أصغر من قيمة (t) الجدولية نقبل بــأن المتغيــرين (X_i) و (X_j) غير مسؤولان أو أنهما لا يسببان الإرتباط الخطي المتعدد طالمــا كان معامل إرتباطهما الجزئي ليس ذات معنوية إحصائية عالية .

مثال : لتطبيق الإختبارات المذكورة آنفاً على الإرتبط الخطي المتعدد ، نعطي نتائج دالة الكلفة محسوبة من قبل فرار وكلاوبر لصيانة البواخر ، وقــد استخدم في تقدير الدالة عينة من المقطع العرضي (a cross-section) لــ(96) باخرة بصيغة :

$$\text{Log } Y = b_0+b_1\log_{X1} + b_2\log_{X2} + \ldots\ldots\ldots b_7\log_{X7} + \log u$$

عندما يكون :

Y = كلفة صيانة الباخرة الواحدة في السنة (بآلاف الدنانير أو الدولارات)

X_1 = عمر الباخرة بالسنوات .

X_2 = حجم الباخرة (الإزاحة بآلاف الأطنان) .

X_3 = الزمن بين الصيانة الكلية .

X_4 = إستهلاك الوقود .

X_5 = متغير أصم (نوعية الوقود مثلاً ديزل أو بخار أو نووي) .

X_5 = 1 اذا كانت الباخرة تشتغل بالديزل

X_5 = 0 اذا لم تكن الباخرة تشتغل بالديزل

X_6 = متغير أصم (Dummy Variable)

X_6 = 1 اذا كانت الباخرة مزودة بالرادار

X_6 = 0 اذا كانت الباخرة غير مزودة بالرادار

X_7 = متغير أصم لوصف في ما اذا كانت الباخرة قد خضعت الى (FRAM) (Fleet Rehabilitation And Modernization) .

X_7 = 1 اذا كان هناك (FRAM)

X_7 = 0 اذا لم يكن هناك (FRAM)

إن النتائج التي تم الحصول عليها من طريقة المربعات الصغرى التقليدية هي :

$$R^2Y._{X1 \ldots\ldots\ldots X7} = 0.80$$

قيمة (F) = 56 مع درجات حرية : $(\upsilon_1=k-1=7)$ ، $(\upsilon2=n-k=96-8=88)$

	$\hat{b}_7$	$\hat{b}_6$	$\hat{b}_5$	$\hat{b}_4$	$\hat{b}_3$	$\hat{b}_2$	$\hat{b}_1$
	-0.16	0.11	-0.03	0.05	-0.79	0.40	0.34
$\delta(\hat{b}_i)$	(0.06)	(0.04)	(0.09)	(0.10)	(0.09)	(0.08)	(0.03)

إن إختبار (F) الكلي والأخطاء المعيارية الصغيرة لمعظم التقديرات يقترح أن الإنحدار ذا معنى بصيغة أن (Y) لا يعتمد فعلياً على منظومة من المتغيرات ($X_7 \ldots X_2, X_1$) ، لإختبار درجة الإرتباط الخطي المتعدد فإن (χ^2*) كان قد حسب ووجد أنه يساوي (261) ، إن القيمة النظرية لـ (χ^2) عند مستوى معنوية (0.05) مع درجات حرية تساوي $\left[\frac{1}{2}k(k-1) = 21\right]$ تساوي (32.7) ، طالما أن (χ^2*) هي أكثر كثيراً من ($\chi^2_{0.05}$) نخلص الى القول أن هناك درجة مهمة من الإرتباط الخطي المتعدد في الدالة .

ولإيجاد أي من المتغيرات يمكن إعتبارها متغيرات تسبب الإرتباط الخطي المتعدد ، نقوم بحساب معاملات الإرتباط الخطي المتعدد وقيم (F) التي ترافقها ضمن منظومة المتغيرات التوضيحية :

$$R^2_{X1 \cdot X2X3 \ldots\ldots X7} = 0.30 \qquad F_{X1} = 6.3$$
$$R^2_{X2 \cdot X1X3 \ldots\ldots X7} = 0.57 \qquad F_{X2} = 19.9$$
$$R^2_{X3 \cdot X1X2 \ldots\ldots X7} = 0.30 \qquad F_{X3} = 6.3$$
$$R^2_{X4 \cdot X1X2 \ldots\ldots X7} = 0.76 \qquad F_{X4} = 47.5$$
$$R^2_{X5 \cdot X1X2 \ldots\ldots X7} = 0.76 \qquad F_{X5} = 47.1$$
$$R^2_{X6 \cdot X1X2 \ldots\ldots X7} = 0.46 \qquad F_{X6} = 12.7$$
$$R^2_{X7 \cdot X1X2 \ldots\ldots X7} = 0.24 \qquad F_{X7} = 4.8$$

من النتائج المذكورة آنفاً يبدو أن العناصر الأكثر تأثراً بالإرتباط الخطي المتعدد هي المتغير (X_4) (إستهلاك الوقود) والمتغير (X_5) (التشغيل بالديزل) .

وأخيراً من أجل إيجاد أي المتغيرات مسؤول عن الإرتباط الخطي المتعدد بين المتغيرين (X_5, X_4) نقوم بحساب معامل الإرتباط الجزئي للمتغيرات التوضيحية وكذلك قيم إحصائية (t) التي ترافق كل معامل من معاملات الإرتباط الجزئي .

إن الحسابات تبدو في الجدول (9−3) وتبين أن سبب الإرتباط الخطـــي المتعدد يعــود علـــى نحـــو رئـــيس الـــى التـــداخل أو الإرتبـــاط المتبـــادل (intercorrelation) بين :

أ− بين (X_4) و (X_5)

ب− بين (X_5) و (X_6)

جدول (9−3)

	X_1	X_2	X_3	X_4	X_5	X_6	X_7
العمر X_1							
الحجم X_2	**r12=0.13 (1.27)**						
الدائرة X_3	**r13=0.21 (2.01)**	**r23=0.27 (2.60)**					
الوقود X_4	**r14=-0.35 (−3.57)**	**r24=0.34 (3.44)**	**R34=0.09 (0.82)**				
ديزل X_5	**r15=-0.27 (1.27)**	**r24=0.21 (-2.03)**	**R35=0.12 (1.15)**	**R45=-0.68 (-8.77)**			
رادار X_6	**r16=0.45 (4.72)**	**r26=0.08 (0.78)**	**R36=-0.35 (-3.51)**	**R46=0.31 (3.13)**	**R56=0.50 (5.51)**		
X_7FRAM	**r17=0.34 (3.38)**	**r27=-0.13 (-1.27)**	**R37=0.06 (0.59)**	**R47=0.40 (4.06)**	**R57=0.27 (2.68)**	**R67=-0.26 (-2.55)**	

*** ملحوظة : الأرقام بين الأقواس تمثل قيم إختبار (t)**

5−32109 بعض الحلول لمشكلة الإرتباط الخطي المتعدد

اذا كانت مشكلة الإرتباط الخطي المتعدد موجودة في الدالة فإن الحلول التي يتم تبنيها تختلف إعتماداً على قساوة مشكلة الإرتباط الخطي المتعدد وعلى وفرة مصادر أخرى للبيانات (مثلاً عينة أكبر أو عينـــات المقطـــع العرضـــي

...الخ) وعلى أهمية العناصر الداخلة في الإرتباط الخطي المتعدد وعلى الغرض الذي قدرت من أجله الدالة وإعتبارات أخرى .

1- إقترح بعض الكتاب أنه اذا لم تؤثر مشكلة الإرتباط الخطي المتعدد على تقديرات المعلمات فإننا يمكن أن نتحمل وجودها في الدالة على الرغم من أن سلامة أو وحدة تقديرات المربعات الصغرى الى مدى معين تضعف .

2- كتاب آخرون يقترحون أنه اذا أثرت مشكلة الإرتباط الخطي على بعض العناصر غير المهمة فبإمكان الفرد أن لا يدخل تلك العناصر في الدالة .

3- إن مشكلة الإرتباط الخطي ربما تؤثر في جزء من مجموعة معلمات ($\hat{b}_s$) بينما بقية المقدرات ربما تبقى نوعاً ما مستقرة ويمكن الإعتماد عليها.

وفي هذه الحالة :

أ- إن المعلمات المقدرة ($\hat{b}'_s$) التي يعتمد عليها ربما تستخدم لأي غرض مثل التنبوء (Forecasting) أو تشكيل السياسة التي تتطلب معلومات يعتمد عليها حول المعلمات الهيكلية (Structural Parameters) .

ب- إن كل التقديرات ربما تستخدم لغرض التنبوء وذلك لأن نمط الإرتباط الخطي المتعدد سوف يستمر بالوجود في فترة التنبوء .

ولكن اذا كانت مشكلة الإرتباط الخطي المتعدد تؤثر على نحو خطير على المعلمات المقدرة للعناصر المهمة في الدالة فعلى الباحث أن يتبنى واحداً من الحلول التصحيحية الآتية :

1- تطبيق طرق تستخدم معلومات كمية خارجية ، ومن أهم هذه الطرق:

أ- طريقة المربعات الصغرى المقيدة (Restricted Least Squares) .

ب- طريقة جمع بيانات المقطع العرضي وبيانات السلاسل الزمنية ، والتي هي حالة خاصة في طريقة المربعات الصغرى المقيدة .

ج- صيغة دوربين (Durbin) للمربعات الصغرى المعممة (Generalised Least Squares) .

د- أسلوب التقدير المختلط (Mixed Estimation Technique) الذي اقترح من قبل ثيل (Theil) وكولد بركر (Goldberger) .

2- زيادة حجم العينة .

من الأمور التي اقترحت حديثاً أن الإرتباط الخطي المتعدد ربما يمكن تجنبه أو تقليل أثره من خلال زيادة حجم العينة بجمع مشاهدات أكثر ، وهكذا يقول كرست (Christ) بزيادة حجم العينة ، فإن التباين المشترك (Covariance) العالي بين المعلمات المقدرة الناتج من وجود الإرتباط الخطي المتعدد في المعادلة يقل أو يتقلص وذلك لأن ذلك التباين المشترك يتناسب عكسياً مع حجم العينة ، وإن هذا صحيحاً أو حقيقياً فقط اذا كان الإرتباط الخطي المتعدد بسبب أخطاء القياس كما هو الحال عندما يكون الإرتباط المتداخل موجوداً في العينة الأصلية فقط وليس في المجتمع الإحصائي لـ(X_s).

فإذا كانت المجتمعات الإحصائية للمتغيرات مرتبطة إرتباطاً خطياً متعدداً وواضحاً فإن زيادة حجم العينة سوف لن يساعد في تقليل أو تقليص علاقات الإرتباط الخطي المتعدد بين المتغيرات .

3- إحلال متغيرات مستقلة محل المتغيرات المتخلفة زمنياً (Lagged Variables) في نماذج التخلف الموزع (Distributed - Lag) .

في الوقت المعاصر هناك الكثير من البحث في الإقتصاد القياسي وجهت البحوث فيه تجاه إستخدام القيم المختلفة زمنياً (lagged Values) للمتغيرات التوضيحية ، وهذا يعني أن الباحثين قد تعرفوا على حقيقة مفادها أن نمطاً معيناً من السلوك يتقرر ليس فقط بالقيم الحالية للمتغيرات التوضيحية ولكن أيضاً بالقيم الماضية (Past Values) لتلك المتغيرات ، على سبيل المثال أنماط الإستهلاك للأفراد تعتمد على الدخول الحالية (Current income) كما هي تعتمد على الدخول الماضية ، ولكن مستويات الدخل الحديثة أو الأكثر حداثة تمارس تأثيراً على قرارات الإستهلاك أكبر من تأثير المستويات الماضية للدخل ، وهكذا فإن الدالة الأصلية ربما تكتب كالآتي :

$$\mathbf{Y_t = b_0 + b_1X_t + b_2X_{(t-1)} + b_3X_{(t-2)} + \ldots\ldots + u_t}$$

ولأسباب واضحة فإن القيم المتتالية X_{t-2},X_{t-1},X_t ... الخ لأي متغير توضيحي (X_i) غالباً مرتبطة ببعضها على نحو قوي جداً ، يمكن تجنب الإرتباط الخطي المتعدد في هذه الحالة من خلال تبني مقترح كويك (Koyck) في إحلال القيم المتخلفة زمنياً (lagged Values) للمتغير التوضيحي (X_i) :

$$\mathbf{Y_t = b_0 + b_1X_t + PY_{t-1} + (u_t - Pu_{t-1})}$$

في هذا النموذج وضعنا بدلاً من كل القيم المتخلفة زمنياً لــ(X) كل من (X_t) (Y_{t-1}) في الدالة ، حيث من المتوقع أن يكون هذين المتغيرين أقل إرتباطاً ببعضها من القيم المتخلفة لــ(X) .

4- إدخال معادلات إضافية في النموذج .

إن مشكلة الإرتباط الخطي ربما تعالج اذا قمنا بإدخال معادلات إضافية في النموذج للتعبير عن علاقات ذات معنى بين المتغيرات ذات الإرتباط الخطي المتعدد (X'_s) . عندما ننظر الى منظومة من المتغيرات المستقلة بإمكان الفرد أو الباحث في معظم الحالان أن يجد علاقات بين (X'_s) وبقية المتغيرات الجديدة التي تجعل للعلاقة معنى إقتصادي ، وبصيغة مباشرة نقول أن الباحث

يعرض تلك العلاقات بصيغة معادلات آنية ، أو نماذج المعادلات الآنية اذا تـــم تشخيصها يمكن أن تقدر على وفـــق المعـــادلات الآنيـــة والشـــكل المختـــزل (reduce form) بوصفها احدى طرق المعادلات الآنية سوف تتجاوز مشـــكلة الإرتباط الخطي فإذا كان النموذج الجديد نموذجاً فوق التشخيص ، فإن الباحـــث ربما يستخدم معلومات إضافية لبعض المعلمات من أجـــل الوصـــول الـــى أو الحصوكل على قيم وحيدة لبقية المعلمات من الشكل المختزل .

الفصل العاشر

مشكلة الإرتباط الذاتي

مشكلة الإرتباط الذاتي

10-1 مقدمـــــــة

من الإفتراضات المهمة في النموذج الخطي الكلاسيكي أن لا يكون هناك إرتباط ذاتي (Autocorrelation) أو إرتباط متسلسل (Serial Correlation) بين قيمة المتغير العشوائي في الفترة الزمنية الحالية وقيمة المتغير العشوائي في الفترة الزمنية اللاحقة لها ، أو نقول ليس هناك إرتباط بين حدود حد الإضطراب (Disturbance Term)(u_i) الداخل في دالة الإنحدار للمجتمع الإحصائي .

10-2 طبيعة مشكلة الإرتباط الذاتي

إن مصطلح الإرتباط الذاتي (Autocorrelation) يمكـن أن يعـرف بوصفه الإرتباط بين أعضـاء سلاسـل مـن المشـاهدات المنظمـة زمنيـاً (كما هي في بيانات السلاسل الزمنية) أو أعضاء المكان في حالة بيانات المقطع العرضي ، وفي سياق الإنحدار فإن نموذج الإنحدار الخطي الكلاسيكي يفترض أن مثل ذلك الإرتباط الذاتي غير موجود فـي حـدود الإضـطراب وبصـيغة المعادلات :

$$E : (u_i u_j) = 0 \qquad i \neq j$$

ببساطة فإن النموذج الكلاسيكي يفترض أن حد الإضطراب لأي مشاهدة لا يتأثر بحد الإضطراب لأي مشاهدة أخرى ، مثلاً اذا كنا نتعامل بسلسلة زمنية فصلية تتضمن إنحدار الإنتاج على رأس المـال والعمـل بوصـفها مـدخلات (Inputs) ونفترض أن هناك إضراب للعمال يؤثر في الإنتاج في فصل واحـد من فصول السنة وأنه ليس هناك سبب يدعو الى الإعتقاد أن هـذا الإضـراب سوف يؤثر في الإنتاج في الفصل التالي من السنة ، وهذا يعني اذا كان الإنتاج منخفضاً في هذا الفصل فليس هناك سبباً يدعو الى التوقع بأن الإنتاج في الفصل التالي سيكون أقل .

وبالطريقة ذاتها اذا كنا نتعامل مع بيانات المقطع العرضي التي تتضمن إنحدار المصروف الإستهلاكي للعائلة على دخل العائلة ، فإن تأثير الزيادة في دخل العائلة على إنفاقها الإستهلاكي فلا نتوقع أن ذلك يؤثر في الإنفاق الإستهلاكي لعائلة أخرى ، ولكن على كل حال اذا كان هناك إعتماد من هذا النوع فإننا لدينا إرتباط ذاتي ، وبصيغة الإشارات :

$$E : (u_i u_j) = 0 \qquad i \neq j$$

في هذه الحالة فإن الإضطراب الذي سببه الإضراب في الفصل الحالي ربما يؤثر في الإنتاج في الفصل التالي أو القادم أو إن الزيادات في الإنفاق من قبل إحدى العوائل يؤثر في زيادة الإنفاق من قبل عائلة أخرى . لعل من المفيد أن ننظر الى بعض الأنماط من حالات الأرتباط الذاتي وحالات عدم وجود الإرتباط الذاتي المعروضة في الشكل (10−1)(e,d,c,b,a) الذي يبين أن هناك أنماط بين (u_i) يمكن أن نميزها أو نراها .

الشكل (10−1−a) يبين النمط الدائري والشكل (10-1-b) و(10-1-c) يؤشران إتجاه خطي صاعد أو نازل لحدود الإضطراب بينما الشكل (10-1-d) يؤشر الإتجاه الخطي والتربيعي في حدود الإضطراب . إن الشكل (10-1-e) هو الوحيد الذي يبين عدم وجود نمط منتظم بين حدود الإضطراب وهنا يدعم إفتراض عدم وجود الإرتباط الذاتي للنموذج الكلاسيكي في الإنحدار الخطي .والآن السؤال الطبيعي هو لماذا يحدث الإرتباط الذاتي أو المتسلسل ؟

ثمة أسباب عدة بعضها كما يأتي :

1− القصور الذاتي (lenteria)

من الخصائص الملحوظة لمهظم السلاسل الزمنية الإقتصادية هي القصور الذاتي ، وكما هو معروف جيداً أن سلاسل زمنية مثل (GNP) أو الأرقام القياسية أو الإنتاج أو الإستخدام أو البطالة تتعرض للدورات (التجارية) وعندما نبدأ من قعر الركود الإقتصادي وعندما يبدأ الإنتعاش الإقتصادي فإن

معظم هذه السلاسل الزمنية تبدأ بالحركة الى الأعلى في هـذا التذبـذب الـى الأعلى، فإن قيمة السلسلة في نقطة احدة أكبر من قيمته في النقطة السابقة .

وهكذا فإن هناك ثمة قوة دافعة أو زخم مبني في تلك السلاسل وتستمر تلك القوة الدافعة الى ان يحصل شيئاً (مثلاً زيادة في سعر الفائدة أو زيادة فـي الضرائب أو كليهما) ليبطيء أو ليقلل السلاسل أو يحركها الى الإنخفاض ، لذلك في الإنحدارات التي تتضمن بيانات سلاسل زمنية تكون المشـاهدات المتتابعـة تدخل في علاقات متداخلة .

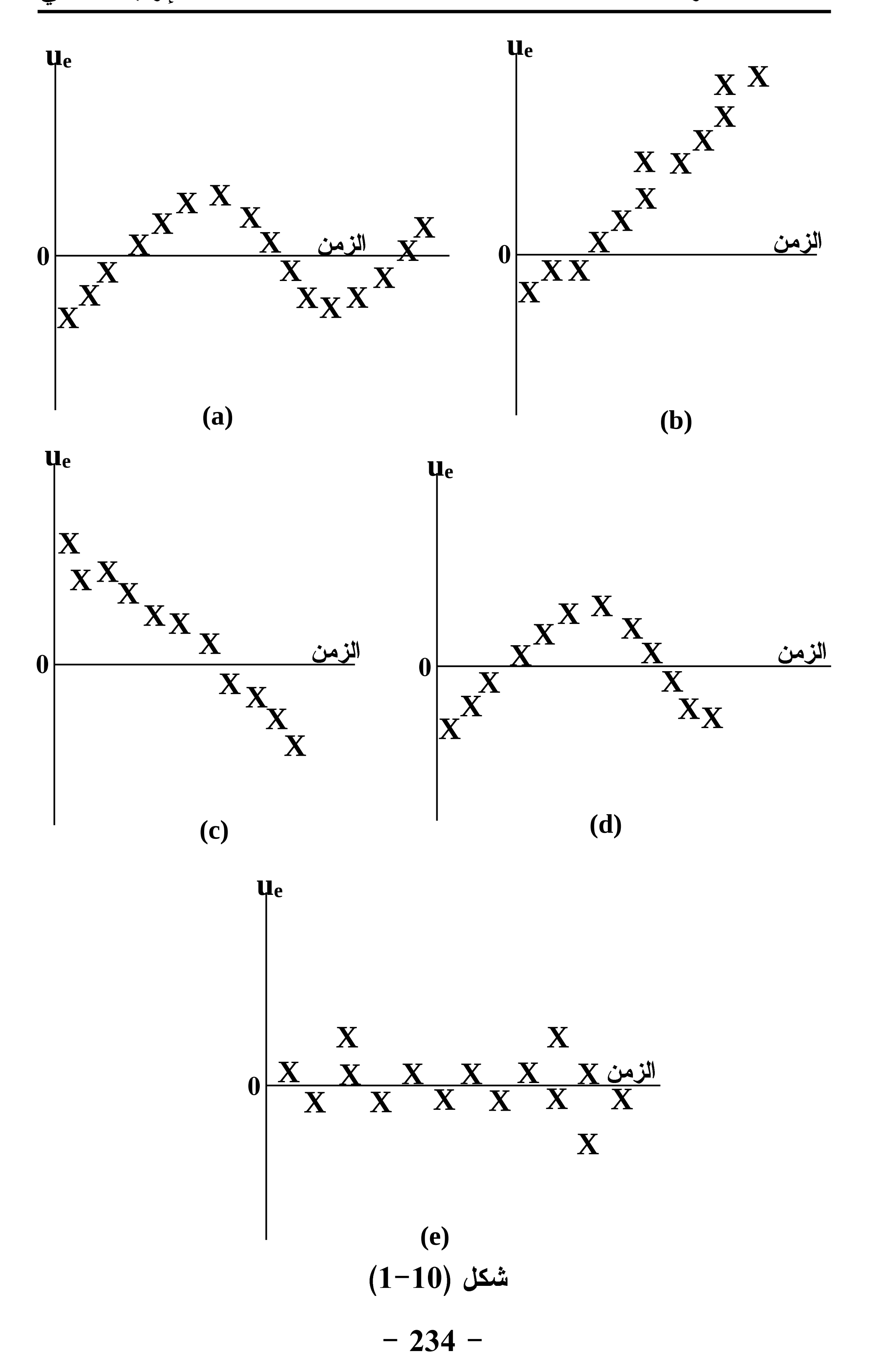
2- تحيز التحديد أو التوصيف (Specification Bias)

حالة إستبعاد متغيرات .

في التحليل التجريبي (Empirical Analysis) عادة يبدأ الباحث مـع نموذج إنحدار ممكن ولكن ربما يكون غير تام ، بعد تحليل الإنحدار فأن الباحث يقوم بفحص في ما اذا كانت نتائج النموذج تتطابق مع التوقعات المسبقة ، فـإذا لم تتطابق فإن عملية معينة يجب أن تجرى على النموذج ، مـثلاً ربمـا يقـوم الباحث برسم البواقي (e_i)(Residuals) التي يتم الحصول عليها مـن تحليـل الإنحدار ، وربما يلاحظ الباحث مثل تلك الأنماط المرسومة في الشكل (10-1) إن هذه البواقي (وهي تقريبات proxies للمتغير العشوائي u_i) ربما تقترح إن بعض المتغيرات التي كانت أصلاً مرشحة ان تكون ضمن الدالة ولكنها لم تدخل في الدالة أو النموذج لأسباب مختلفة يجب أن تدخل في النموذج.

هذه هي حالة أبعاد المتغيرات التي تبين تحيزاً في التوصيف للنموذج ، غالباً ما يحدث إن إدخال مثل تلك المتغيرات في الدالة يحـذف وجـود نمـط الإرتباط المشاهد بين البواقي مثلاً نفترض أن لدينا نموذج الطلب الآتي :

$$\mathbf{Y_t = b_1 + b_2X_{2t} + b_3X_{3t} + b_4X_4t + u_t \ldots\ldots(1\text{-}10)}$$



شكل (1−10)

عندما يكون :

Y = كمية لحم البقر المطلوب

X_2 = سعر لحم البقر

X_3 = دخل المستهلك

X_4 = سعر لحم الخنزير

t = الزمن

ولكن لبعض الأسباب نحدد الإنحدار التالي :

$$Y_t = b_1 + b_2X_{2t} + b_3X_{3t} + b_4X_4t + \upsilon_t \;......(2\text{-}10)$$

الآن اذا كانت المعادلة (10−1) هي النموذج الصحيح او الحقيقي أو المعبر عن العلاقة الحقيقية ، فإن إستخدام المعادلة (10−2) معادل أو مساوي لجعل :

$$\upsilon_t = b_4X_4t + u_t \;...........(3\text{-}10)$$

والى مدى أن سعر لحم الخنزير يؤثر في إستهلاك لحم البقر فإن حد الخطأ أو حد الإضطراب (υ) سوف يعكس نمطاً أو نموذجاً منتظماً وهكذا خلق إرتباط ذاتي خاطيء ، والإختبار البسيط لهذا هو أن نستخدم النموذج (10−1) والنموذج (10−2) ونرى في ما اذا كان هناك إرتباطاً ذاتياً مشاهداً في النموذج (10−3) يختفي عندما نستخدم النموذج (10−1) .

3− الشكل الدالي غير الصحيح (Incorrect Functional Form)

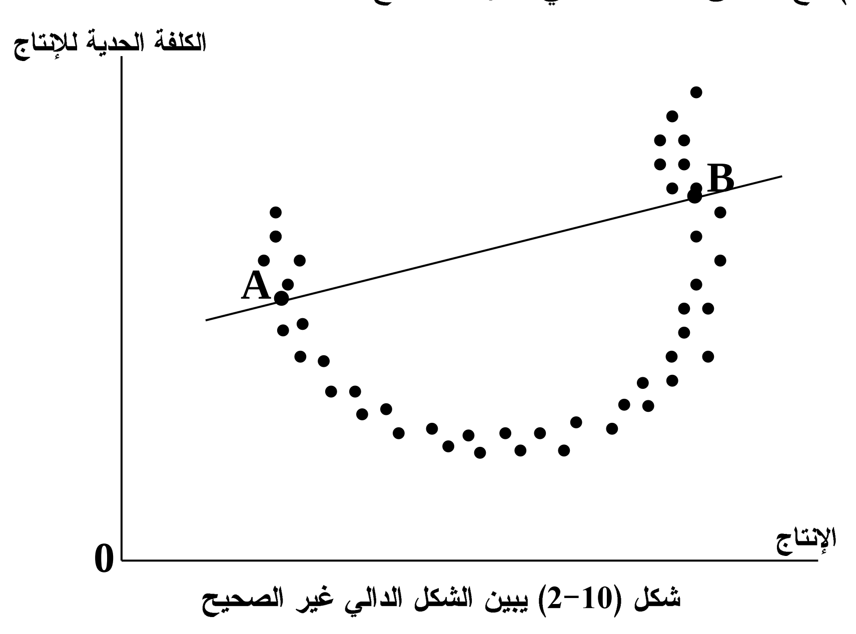
لنفترض أن النموذج الحقيقي أو الصحيح في دراسة الكلفة – الإنتاج كما ياتي :

$$\text{الكلفة الحدية} = b_1 + b_2\,(\text{الإنتاج}) + b_3\,(\text{الإنتاج})^2 + u_i \;.......... (4\text{−}10)$$

ولكن اذا قدرنا النموذج الآتي :

$$\text{الكلفة الحدية} = {}_2\alpha\,{}_1 + \alpha\,(\text{الإنتاج}) + \upsilon i \;....... (5\text{−}10)$$

إن منحنى الكلفة الحدية الذي يرافق النموذج الحقيقي يبدو في الشكل (10−2) مع منحنى الكلفة الخطي غير الصحيح :



شكل (10−2) يبين الشكل الدالي غير الصحيح

وكما يبدو في الشكل (10−2) فإن المسافة بين النقطتين (A)(B) تبين منحنى الكلفة الحدية الخطي الذي سوف يبين تقديراً للكلفة الحدية يفوق الكلفة الحدية الحقيقية ، بينما ما بعد النقطتين فإن المنحنى الخطي سوف يعطي تقديراً للكلفة الحدية أقل من الكلفة الحدية الحقيقية ، وهذا متوقع لأن حد الإضطراب (υi) الذي يساوي في الواقع القيمة الآتية :

(الإنتاج) $^{2} + u_i$

ولذلك فهو يمسك أو يعبر عن التأثير المنتظم لتربيع الإنتاج (**(الإنتاج)** ^{2}) في الكلفة الحدية ، وفي هذه الحالة فإن (υi) سوف يعكس الإرتباط الذاتي بسبب إستخدام الشكل الدالي غير الصحيح .

4- الظاهرة العنكبوتية (Cobweb Phenomenon)

إن العرض من عدة سلع زراعية يعكس ما يدعى بالظاهرة العنكبوتية ، والتي يكون فيها رد الفعل من قبل العرض الى السعر مع تخلف زمنـــي لفتــرة واحدة ، وذلك لأن قرارات العرض تأخذ وقتاً للتنفيذ (فترة النضوج) ، وهكـــذا في بداية فترة زراعة المحاصيل لهذه السنة فإن المزارعين متأثرون بالأســعار السائدة في السنة السابقة بحيث تكون دالة العرض هي :

$$\mathbf{Q_s = b_0 + b_1P_{t-1} + u_t}$$

ولنفترض أنه في نهاية الفترة (t) وجدنا ان سعر السنة (t) (P_t) ظهــر أنه أقل من (P_{t-1}) ، لذلك في الفترة (t+1) فإن المزارعين ربمـــا يقـــررون أن ينتجوا أقل مما أنتجوا في الفترة (t) .

اذاً من الواضح في هذه الحالة أن قيم المتغير العشوائي (u_i) من غيـــر المتوقع أن تكون عشوائية وذلك لأنه اذا أنتج المزارعون كميات أكثر في السنة (t) مما انتجوا في السنة (t-1) ، فإنهم سوف يقللون إنتاجهم في الســـنة (t+1) وهكذا مما يقود الى النموذج العنكبوتي .

5- التخلف الزمني (Time Lags)

في تحليل إنحدار السلاسل الزمنية للمصروفات الإستهلاكية على الدخل من المعتاد أن نجد أن المصروفات الإستهلاكية في الفترة الحالية تعتمد من بين عناصر أخرى على المصروفات الإستهلاكية في الفترة السابقة بمعنى :

الإستهلاك في السنة t = $b_1 + b_2$ الدخل في السنة t + b_3 الإستهلاك في السنة $t-1$ + ...u_t (6-10)

إن إنحداراً مثل الموجود في المعادلة (10-6) يعرف بأنه إنحداراً ذاتياً (Autoregression) وذلك أحد المتغيرات التوضيحية هو القيمة المتخلفة زمنياً (Lagged Value) للمتغير المعتمد ، إن المنطق الذي يدعم هذا النموذج هو أن المستهلكين لا يغيرون عاداتهم الإستهلاكية في الغالب لأسباب نفســية وتقنيــة ومؤسسية ، فإذا أهملنا الحد المتخلف زمنياً في النموذج (10-3) فإن حد الخطأ

الناتج سوف يعكس نمطاً منتظماً بسبب تأثير الإستهلاك المتخلف على الإستهلاك الحالي (Current Conoumption) .

6- معالجة البيانات (Manipulation of Data)

إن البيانات الخام عادة في التحليل التجريبي تتم معالجتها ، مثلاً في تحليلات الإنحدار للسلاسل الزمنية (Time-Series) التي تتضمن بيانات فصلية، وإن مثل هذه البيانات عادة تشتق من بيانات شهرية من خلال إضافة بيانات كل ثلاثة أشهر وتقسيم الناتج على ثلاثة ، إن هذه المعدلات تدخل تمهيداً (Smothness) في البيانات من خلال طمس التقلبات (Fluctuations) في البيانات الشهرية ، في البيانات الفصلية تبدو أكثر إعتدالاً أو نعومة أو تمهيداً من البيانات الشهرية ، وإن هذه النعومة أو الإعتدال ربما يقود الى نمط منتظم في حدود الإضطراب وهذا يتضمن إدخالاً للإرتباط الذاتي ، ثمة مصدر آخر لمعالجة البيانات وهو تحريف البيانات أو تقدير البيانات على أساس بيانات متوافرة (Extrapolation of data) مثلاً تعداد السكان يجري كل عشرة سنوات في بلد معين بحيث مثلاً يكون التعداد الأخير 1987 والسابق له هو تعداد 1977 والأسبق هو تعداد 1967 وقبله تعداد 1957 وهكذا ، والآن اذا كانت هناك حاجة للحصول على بيانات لعدد أو لبعض السنوات ضمن فترة متداخلة بين تعدادين مثلاً (1977-1987) فإن العملية الشائعة أن تحرف البيانات على أساس إفتراضات لأغراض محددة .

10-3 النتائج المترتبة على وجود مشكلة الإرتباط الذاتي (Consequance of Autocorrelation)

عندما يعاني حد الإضطراب (disturbance Term) من الإرتباط المتسلسل أو الذاتي بين قيمه ، تحصل النتائج الآتية في حالة تطبيقنا لطريقة المربعات الصغرى الإعتيادية (OLS) :

1- إن قيمة المعلمات المقدرة تتأثر بوجود الإرتباط الذاتي .
2- إن الأخطاء المعيارية للمعلمات المقدرة تتأثر أيضاً بوجود الإرتباط الذاتي .
3- ولكن حتى لو كان الإرتباط الذاني موجوداً بين قيم المتغير العشوائي ، فإن تقديرات طريقة المربعات الصغرى الإعتيادية (OLS) للمعلمات لا تكون متميزة إحصائياً ، بمعنى إن القيمة المتوقعة تساوي المعلمـــة الحقيقيـــة (True parameter) .
4- في حالة وجود الإرتباط الذاتي في قيم حد الإضطراب فإن التباين المقـــدر بطريقة المربعات الصغرى (OLS) للمعلمات المقدرة يكون أكبـــر مـــن التباين المقدر بالطرق الإقتصادية القياسية الأخرى .
5- في حالة وجود الإرتباط الذاتي في قيم حد الإضطراب فإن تباين المتغيـــر العشوائي (u_i) يكون تقديره قليلاً (Underestimated) على نحو خطير وهذه الخطورة تكون أكبر في حالة وجود إرتباط ذاتي موجب .
6- اذا كانت قيم المتغير العشوائي مرتبطة ذاتياً فإن التنبوءات (Predictions) الحاصلة على أساس تقديرات المربعات الصغرى الإعتيادية سوف تكـــون غير كفوءة بمعنى أنها تمتلك تباين أكبر من ذلك التباين المقـــدر بطـــرق قياسية أخرى .

10-4 إختبارات الإرتباط الذاتي

من الأمور المعروفة أن وجود الإرتباط الذاتي ونمطه يمكن أن نصـــل اليه أو نحصل عليه من خلال رسم (plotting) بواقي الإنحدار ، أما مقابل قيمه المتخلفة أو الزمن .

ولكن هناك ثمة إختبارات أكثر دقة للكشف عن مشكلة الإرتباط الـــذاتي (Autocorrelation) ، من الإختبارات المطبقة تقليدياً إختبار معدل فون نيومن (Von Veumann) وإختبار دربن-واطسون (Durbin-Watson) :

وهذا الإختبار يمكن أن يأخذ الصيغة الآتية :

$$\frac{\delta^2}{s_x^2}=\frac{\sum_{t=2}^{n}\left(X_t-X_{t-1}\right)^2/(n-1)}{\sum\left(X-\overline{X}\right)^2/n}$$

وهذا هو نسبة تباين الفروقات الأولى (First-differences) لأي متغير مثل (X) على تباين (X) . إن معدل فون نيومن يمكن تطبيقه للسلاسل الزمنية المشاهدة مباشرة وللمتغيرات العشوائية وهي المتغيرات التـي تكـون قيمتها المتتابعة غير مرتبطة ذاتياً ، في حالة المتغير العشوائي (u) فإن قيمه غير قابلة للمشاهدة المباشرة ، ولكنها مقدرة عن طريق قيم البواقي الحاصلة من طريقـة المربعات الصغرى الإعتيادية (OLS) (e'_s) ، وللعينات الكبيرة التـي تكـون مشاهداتها أكبر من ثلاثين مشاهدة (n أكبر من 30) فإن معدل فون نيومن ربما يكون كما يأتي :

$$\frac{\delta^2}{S_e}=\frac{\sum_{i=2}^{n}\left(\ell_t-\ell_{t-1}\right)^2/(n-1)}{\sum\left(\ell_t-\overline{\ell}\right)^2/n}$$

ويمكن ان يطبق تقريباً مع ($\overline{\ell}=0$) بالتعريف ، ولكن ذلك غير ممكـن لأن قيم بواقي (ℓ) طريقة المربعات الصغرى الإعتيادية (OLS) موزعة على نحو مستقل . لذلك فإن إختبار فون نيومن لا يمكن تطبيقه لإختبـار الإرتبـاط الذاتي لقيم المتغير العشوائي بخاصة اذا كانت العينة صغيرة أي أقل من ثلاثين مشاهدة (n أقل من ثلاثين) .

أما دربن وواطسن (Durbin-Watson) فقد إقترحا إختباراً يطبق في حالة العينات الصغيرة ، ولكن ذلك الإختبار ملائم فقط للإرتبـاط الـذاتي مـن الدرجة الأولى (The First- Order autoregressive) ($u_t=\rho u_{t-1}+\upsilon_t$)
وهذا الإختبار يأخذ الخطوات الآتية :

فرضية العدم $H_0 : \rho = 0$

أو أن (H_0) : أن قيم (u) غير مرتبطة ببعضها ذاتياً إرتباطاً من الدرجة الأولى ، إن هذه الفرضية تختبر ضد :

الفرضية البديلة $H1 : \rho \neq 0$

أو أن (H_1) : أن قيم (u) مرتبطة ببعضها ذاتياً إرتباطاً من الدرجة الأولى .

لإختبار فرضية العدم نستخدم إحصاءة دربن–واطسون (Durbin-Watson statistic) :

$$d = \frac{\sum_{t=2}^{n}\left(\ell_t - \ell_{t-1}\right)^2}{\sum_{t=1}^{n}\ell_t^2}$$

إن قيم (d) تقع بين (0)(4) ، وعندما تكون (d=2) فإن (ρ=0) ، وهكذا فإن إختبار فرضية (H_0: ρ=0) يساوي إختبار فرضية ($H_0 : d = 2$) ، وعندما نوسع إحصاءة (d) نحصل على :

$$d = \frac{\sum_{t=2}^{n}\left(\ell_t - \ell_{t-1}\right)^2}{\sum_{t=1}^{n}\ell_t^2} = \frac{\sum_{t=2}^{n}\left(\ell^2{}_t + \ell^2{}_{t-1} - 2\ell_t\ell_{t-1}\right)}{\sum_{t-1}^{n}\ell^2{}_t}$$

$$d = \frac{\sum_{t=2}^{n}\ell_t^2 + \sum_{t=2}^{n}\ell_{t=1}^2 - 2\sum_{t=2}^{n}\ell_t\ell_{t-1}}{\sum_{t=1}^{n}\ell_t^2}$$

ولكن في العينات الكبيرة فإن الصيغ :

$$\sum_{t=2}^{n}\ell_t^2 \qquad \sum_{t=2}^{n}\ell_{t-1}^2 \qquad \sum_{t=1}^{n}\ell_t$$

تقريباً متساوية ، ولذلك ربما يمكن ان تكتب الإحصاءة :

$$d \approx \frac{2\sum \ell_{t-1}^2}{\sum \ell_{t-1}^2} - \frac{2\sum \ell_t \ell_{t-1}}{\sum \ell_{t-1}^2}$$

$$d \approx \left(1 - \frac{\sum \ell_t \ell_{t-1}}{\sum \ell_{t-1}^2}\right)$$

ولكن :

$$\left(\frac{\sum \ell_t \ell_{t-1}}{\sum \ell_{t-1}^2}\right) = \hat{\rho}$$

ولذلك :

$$d \approx 2(1 - \hat{\rho})$$

ومن هذه الصيغة يبدو واضحاً أن قيم (d) تقع بين (0)(4) .

أولاً : اذا لم يكن هناك إرتباطاً ذاتيــاً ($\hat{\rho} = 0$) وإن (d=2) وهكــذا اذا وجدنا من بيانات العينة أن (d^*) المحسوبة ($d^* \approx 2$) فإننا نقبل أنه لــيس هنـــاك إرتباطاً ذاتياً في الدالة .

ثانياً : اذا كان ($\hat{\rho} = +1$) فإن (d=0) ويكون لدينا إرتباطاً ذاتياً موجباً ، ولذلك اذا كانت ($0 < d^* < 2$) فإن هناك بعض درجات مـــن الإرتبـــاط الـــذاتي الموجب ويكون ذلك الإرتباط قوياً كلما إقتربت (d^*) من الصفر .

ثالثاً : اذا كانت ($\hat{\rho} = -1$) فإن (d=4) وعندئذ يكون لدينا إرتباطاً ذاتياً تاماً وسالباً ، ولذلك اذا كانت ($2 < d^* < 4$) فإن هناك درجة من الإرتباط الذاتي السالب ويكون ذلك الإرتباط أكثر قوة كلما إرتفعت قيمة (d^*) .

يجب أن يكون واضحاً أن إختبار دربن–واطسون يختبر فرضية العـدم في عدم وجود إرتباط ذاتي (ρ=0) بصورة غير مباشرة من خـلال إختبـار فرضية مساوية للأولى تقول أن (d=2) .

الخطوة الأخرى ان يتم إستخدام بواقي العينة (ℓ_t´s) وحسـاب القيمـة التجريبية لإحصاءة دربن–واطسون (d^*) ، وأخيراً فإن القيمة التجريبية لـ(d^*) يجب أن تقارن مع القيم النظرية لـ(d) تلك القيم لـ(d) التي تعرف المنطقـة الحرجة للإختبار ، إن مشكلة هذا الإختبار هي أن التوزيع التام لـ(d) غيـر معروف ، ولكن دربن–واطسون كانا قد وضعا حداً أعلى لـ(d)(d_u) وحداً أدنى (d_L) لمستويات معنوية (d) الملائمة لإختبار فرضية عدم وجود الإرتباط الذاتي من الدرجة الأولى ضد الفرضية البديلة في وجود الإرتباط الذاتي الموجب مـن الدرجة الأولى ، لقد وضع دربن–واطسون جدولاً لتلك القيم العليا والقيم الـدنيا عند مستوى (5%و1%) من المعنوية ، وهذه الجداول يفترض أن تكون فيها قيم (u) إعتيادية ومتجانسة وغير مرتبطة ذاتياً .

إن الإختبار ذاته يقارن القيمة التجريبية (d^*) المحسـوبة مـن بـواقي الإنحدار مع الحد الأدنى (d_L) والحد الأعلى (d_u) في جداول دربن–واطسـون وقيمها المحولة ($4-d_L$) ($4-d_u$) ، إن المقارنـة بإسـتخدام (d_L) (d_u) تبحـث إمكانية وجود إرتباطاً ذاتياً موجباً ومقارنة مع ($4-d_L$) ($4-d_u$) تبحث إمكانيـة وجود إرتباطاً ذاتياً سالباً :

1– اذا كانت (d^*) المحسوبة أقل من (d_L) النظرية فإننا نرفض فرضية العـدم التي تنص على عدم وجود إرتباط ذاتي ، نقبل أن هناك إرتباطاً ذاتياً موجباً من الدرجة الأولى .

2– اذا كانت (d^*) المحسوبة أكبر من ($4-d_L$) فإننا نرفض فرضية العدم التـي تنص على عدم وجود الإرتباط الذاتي ونقبل أن هناك إرتباطاً ذاتياً سـالباً مـن الدرجة الأولى .

3- اذا كانت $(d4<d^*<(4-d4))$ فإننا نقبل فرضية العدم التي تنص على عدم وجود الإرتباط الذاتي .

4- اذا كانت $(d_L<d^*<d4)$ أو اذا كانت $((4-d4)< d^*< (4-d_L))$ فإن الإختبار يكون غير حاسم .

10-5 بعض النواقص في إختبار دربن – واطسون

أولاً : إن الإحصاءة (d) تعد مقياساً غير ملائم للإرتباط الذاتي اذا كان هناك بين المتغيرات التوضيحية قيماً متخلفة زمنياً (Lagged Values) لمتغير داخلي .

ثانياً : إن مدى قيم (d) التي تغطي مساحة منطقة عدم الحسم

$(d_L<d^*<d4)$ و $(4-d4)< d^*< (4-d_L)$ تشكل عثرة في طريق تطبيق إختبار دربن – واطسون .

ثالثاً : إن إختبار دربن–واطسون غير ملائم لإختبار درجات أعلى للإرتباط المتسلسل ، أو لأشكال أخرى من الإرتباط الذاتي مثلاً الإرتباط الذاتي المتسلسل في الأشكال غير الخطية لقيم (u_t).

10-6 الحلول لمشكلة الإرتباط الذاتي

إن الحل الملائم المعتمد او الذي يتم إختياره لحالة إرتباط ذاتي معينة يعتمد أساساً على مصدر (Source) الإرتباط الذاتي ، وهكذا اذا كان مصدر الإرتباط الذاتي هو متغيرات محذوفة فإن الإجراء الملائم هو إدخال تلك المتغيرات ضمن منظومة او مجموعة المتغيرات التوضيحية .

من الممكن أن نوضح حدوث حالة شبه إرتباط ذاتي (quasi-auto correlation) مع التأثير الراجعي في دالة الإستهلاك ، فالإستهلاك في الفترة الزمنية (t) يعتمد ليس فقط على الدخل الحالي

(Current Income) (X_t) وإنما أيضاً على مستوى الخل في الفترات الزمنية السابقة ، يمكن كتابة هذا النمط من دالة الإستهلاك بأبسط شكل من أشكالها كما يأتي :

$$C_t = b_0 + b_1X_t + b_2X_{t-1} + u_t$$

فإذا حذف متغير الدخل المتخلف فترة زمنية واحدة فإن تـأثيره سـوف ينعكس في قيم المتغير العشوائي (u_t) وربما أيضاً في معلمـة متغيـر الـدخل الحالي ، والتي ستكون معلمة متحيزة .

إن الإرتباط الذاتي سوف يحذف أو ينتهي في هذه الحالة مـن خـلال إدخال متغير الدخل المتخلف زمنياً في الدالة بوصفه متغيراً توضيحياً للإستهلاك الحالي (Current Consumption) ، ومن أبسط الطرق لمعرفة في ما اذا كان الإرتباط الذاتي بسبب حذف متغيرات هو أن نقـوم بتحليـل إنحـدار البـواقي (residuals) (ℓ_s) على المتغيرات التي تعد متغيرات توضيحية ملائمة علـى أسس نظرية مسبقة للظاهرة قيد الدراسة .

بالطريقة نفسها اذا كان مصدر الإرتباط هو التحديـد الرياضـي غيـر الملائم للعلاقة ، فإن المعالجة تتم بتغيير الشكل الخطي الأولي وهذا يمكـن أن يبحث بوساطة تحليل إنحدار البواقي على درجـات قـوة أعلـى للمتغيـرات التوضيحية أو الحساب في علاقة خطية باللوغاريتم .

الفصل الحادي عشر

مشكلة عدم ثبات التباين

مشكلة عدم ثبات التباين

11-1 مقدمــــــة (Hetroscedasticity Problem)

من الإفتراضات المهمة المتعلقة بالمتغير العشوائي (u) هو ان إحتمالية توزيع للمتغير العشوائي تبقى نفسها عبر كل مشاهدات (X) ، وبخاصـــة فـــإن تباين كل (u_i) نفسه لكل قيم المتغير المستقل ، وبالرموز لدينا :

$$Var(u) = E\{(u_i - E(u))\}^2$$

$$= E(u_i)^2 = \delta_u^2 \quad \text{ثابت}$$

إن هذا الإفتراض يعرف بوصفه إفتراض الثبات أو التجانس أو إفتراض ثبات تباين الـــ(u_s) ، فإذا لم يتحقق في أية حالة خاصة فإننا نقول أن (u'_s) هي غير متجانسة (Hetroscedasticity) :

$$Var(u_i) = \delta_{ui}^2 \quad \text{غير ثابت}$$

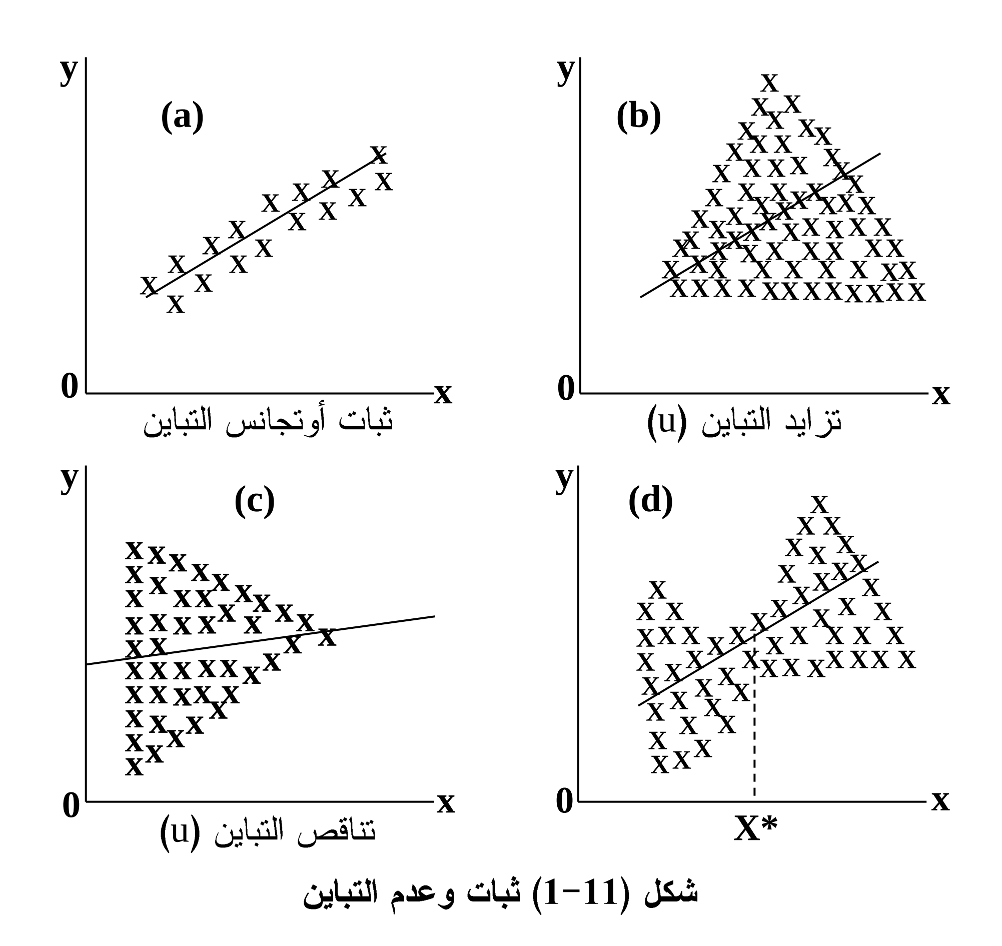
11-2 التفسير والعرض بصيغة شكل الإنتشار لثبات وعدم ثبات التباين

إن معنى إفتراض ثبات التباين هو أن إنحراف (Variation) كل (u_i) حول وسطه الحسابي المساوي للصفر لا يعتمد على قيمة (X) ، إن تباين كـــل (u_i) يبقى نفسه بغض النظر عن القيم الصغيرة والقيم الكبيرة للمتغير المستقل .

إن (δ_u^2) ليس دالة في (X_i) ، وهذا يعني :

$$\delta_u^2 \neq f(X_i)$$

وفي شكل الإنتشار فإن تجانس التباين (Homoscedasticity) يمكـــن أن يوضــــح مــــن الإنتشــــار العشــــوائي (u'_s) ضــــمن مســــافة ثابتــــة (Constant Distance) حول خط الإنحدار :



شكل (11−1) ثبات وعدم التباين

فإذا كان (δ_u^2) غير ثابت ولكن قيمه (its values) تعتمد علـــى قـــيم (X) فإننا ربما نكتب ذلك بالصيغة الآتية :

$$\delta_u^2 = f(X)$$

إن حالة عدم ثبات أو عدم تجانس التبـــاين ((Hetroscedasticity)) تبـــدو أو تتوضح من خلال تزايد أو تناقص إنتشار (dispersion) المشاهدات عن خـــط الإنحدار ، إن نمط المشاهدات على شكل الإنتشار يعتمد على شكل عدم الثبـــات أو عدم التجانس للتباين بمعنى يعتمد على شكل العلاقة بين (δ_u^2) و (X_i) .

ثمة ثلاثة أشكال لحالة عدم تجانس التباين فـــي الشـــكل (11−1) ففـــي الشكل (b 1-11) تم تصوير حالة تزايد تباين الـــ(u'_s) فكلما ازداد (X) يـــزداد

تباين (u) ، وهذا يمثل حالة عدم ثبات أو عدم تجـانس التبـاين شـائعة فـي التطبيقات الإقتصادية القياسية ، في الشكل (c 1-11) نبين نمط تنـاقص عـدم التجانس للتباين ، فكلما أخذ (X) قيماً أكبر فإن إنحراف المشاهدات عـن خـط الإنحدار يتناقص أو يصغر ، بمعنى إن تباين المتغير العشوائي يتغير بالإتجـاه المعاكس مع المتغير التوضيحي ، وأخيراً في الشكل (d 1-11) رسمنا الحالـة الأكثر تعقيداً من حالات عدم تجانس التباين وهي أن تباين المتغير العشوائي (u) يتناقص في البداية كلما أخذ المتغير التوضيحي (X) قيماً أكبـر ، ولكـن بعـد مستوى معين لـ(X) مثلاً (X^*) فإن تباين (u) يزداد مع زيادة المتغيـر (X) ، يجب أن يكون واضحاً إن نمط عدم تجانس التباين يعتمد على إشـارات وقـيم معاملات العلاقة : ($\delta^2 = f(X_i)$) طالما أن (u_i) هي قيم غير ممكن مشـاهدتها فإننا لانعرف النمط الحقيقي لعدم تجانس التباين ، ولكن في البحوث الطبيعية فإن المختصين بالإقتصاد القياسي عادةً يضعون الإفتراض الملائم إن عدم تجـانس التباين هو بشكل :

$$\delta^2_{ui} = K^2 X^2$$

عندما يكون (K) ثابت ويمكن أن يقدر من النموذج .

11-3 إمكانية وجود إفتراض تجانس التباين

في تطبيقات إقتصادية قياسية عدة فإن إفتراض ثبات التبـاين للمتغيـر العشوائي ربما يتوقع ان لا يحدث أو ربما يكون غير موجود ، وهنا يمكن فهمه بسهولة اذا أخذنا بنظر الإعتبار العناصر التي يتم إمتصاص آثارها بوساطة حد الإضطراب (Disturbance Term) أو المتغير العشوائي ، قلنا سـابقاً أن (u) يعبر عن التأثير على المتغير المعتمد من قبل الإخطاء في قياس المتغير المعتمد والمتغيرات المحذوفة ، وعلى كلا الجانبين هناك أسباب للتوقع أن تباين المتغير

العشوائي يختلف عبر الزمن أو يختلف على نحو منتظم مع المتغير التوضيحي (X) في معظم الحالات . وهكذا اذا ازداد (Y) فإن أخطاء القياس (Error of Measurment) تميل الى الزيادة أيضاً وذلك لأنه يصبح من الصعوبة جمع بيانات وتدقيق إتساق تلك البيانات ودرجة الوثوق بها ، والأكثر من ذلك فإن أخطاء القياس تميل الى التراكم عبر الزمن ، بحيث أن حجمها يزداد ، وفي هذه الحالة فإن تباين المتغير العشوائي يزداد مع زيادة قيم المتغير المستقل أو التوضيحي (X) . ومن ناحية أخرى فإن أساليب أخذ العينات والطرق المختلفة في جمع البيانات في تطور مستمر وهكذا فإن أخطاء القياس ربما تتناقص وفي هذه الحالة فإن (δ^2) يتناقص عبر الزمن . والأكثر أهمية من كل هذا هو أن متغيرات محذوفة من الدالة عدة تميل الى التتغيير بالإتجاه نفسه مع (X) المتغير التوضيحي ، وهكذا مسببة زيادة في الإنحراف (Variation) في المشاهدات عن خط الإنحدار .

مثلاً : إفترض أن لدينا عينة مقطع عرضي (A Crossection Sample) لميزانيات العوائل والتي منها نريد أن نقيس دالة الإدخار (Saving Function) :

$$\mathbf{S_i = b_0 + b_1Y_i + U_i}$$

عندما يكون :

S_i= إدخارات العوائل.

Y_i= دخل العوائل .

في هذه الحالة فإن إفتراض ثبات تباين المتغير العشوائي (u) غير ملائم وذلك لإن العوائل ذات المستوى العالي من الدخل تعرض إختلافات أكبر كثيراً في سلوكها الخاص بالإدخارات من العوائل ذات الدخل المنخفض ، فالعوائل ذات الدخل العالي تميل الى الإلتصاق بمستوى معين من المعيشة وعندما ينخفض دخلها تقوم بالإستقطاع من الإدخارات وليس من الإنفاق الإستهلاكي

لهم، بينما العوائل ذات الدخل المنخفض تدخر لأغراض معينة ، وهكذا فإن نمط إدخاراتهم أكثر إنتظاماً . وهذا يتضمن أنه عند المستويات العالية من الدخول ، فإن (u_i) سيكون عالياً بينما عند المستويات المنخفضة للدخل فإن (u_i) سيكون صغيراً ، ولذلك فإن إفتراض ثبات أو تجانس تباين (u_i) سيكون غير موجود ، عندما تقدر دالة الإدخارات من المقاطع العرضية لميزانية الأسرة .

ثمة مثالاً آخر اذا أخذنا عينة مقطع عرضي من شركات في صناعة معينة من أجل إستخدامها لتقدير دالة إنتاج بصيغة كوب – دوغلاس (Cobb-Douglas) :

$$X = b_0 . L^{b1} . K^{b2} \ell^{u}$$

إن (u) في هذه الحالة يمتص تأثير الإختلافات التكنولوجية لمختلف الشركات وكذلك تأثير الإختلافات في المهارات والتنظيم بين الشركات وعناصر أخرى ، إن هذه العناصر لا تتغاير على نحو كبير في الشركات الصغيرة ، بينما من المتوقع أن تتغاير (Vary) في الشركات الكبيرة ، ولذلك فإن قيم (u) سوف لا تتجانس (Hetroscedastic) .

وخلاصة القول ان هناك وعلى أسس مسبقة أسباب للإعتقاد بإن إفتراض تجانس تباين المتغير العشوائي ربما في مناسبات عديدة يتم مخالفته في التطبيق . ولذلك يصبح من المهم جداً أن نتفحص ما يترتب على وجود حالة عدم تجانس أو عدم ثبات تباين المتغير العشوائي وتأثيرها على المعلمات المقدرة وعلى الإنحرافات المعيارية لتلك المعلمات .

11-4 النتائج المترتبة على وجود مشكلة عدم تجانس أو عدم ثبات تباين المتغير العشوائي (u)

اذا لم يتحقق شرط أو إفتراض تجانس تباين المتغير العشوائي فإننا نتوقع أن يكون لدينا المترتبات (Consequences) الآتية :

1- أن لا نستطيع تطبيق قوانين التباينات للمعلمات من أجـــل إجـــراء إختبارات المعنوية الإحصائية وبناء حدود الثقة للمعلمات .

2- اذا كانت قيم (u) غير ثابتة التباين أو غير متجانسة التباين ، فـــإن المعلمات المقدرة بطريقة المربعات الصغرى الإعتيادية ليس لها صفة أقل تباين في صنف المعلمات المقدرة غير المتحيزة بمعنى أن تلك المعلمات غير كفـــوءة في العينات الصغيرة وفي العينات الكبيرة .

3- إن المعلمات المقدرة سوف تبقى غير متحيزة إحصائياً ، بمعنى أنه حتى لو أنه تباين (u) غير ثابت أو غير متجانس فإن المعلمات المقـــدرة ($\hat{b}_i$) سوف لن يكون لها تحيزاً إحصائياً ، أي إن القيمة ســـوف تســـاوي المعلمـــات الحقيقية :

$$E(\hat{b}_i) = b_i$$

4- إن التنبوء بقيم (Y) لقيمة معينة من (X) المعتمدة أو المتأسسة على المعلمات المقدرة ($\hat{b}_s$) من البيانات الأصلية ، وذلك لأن لهـــا تبـــاين عـــالي ، بمعنى ان التنبوء يتضمن تباينات (Variances) المتغير العشـــوائي وتباينـــات المعلمات المقدرة والتي هي ليست صغرى بسبب حدوث عدم تجانس التباين .

11-5 إختبارات ثبات أو عدم تجانس التباين

لقد إقترحت إختبارات مختلفة للكشف عن عدم تجانس التبـــاين ، ثمـــة إختبارين أساسين وهما الأبسط في التطبيق .

11-5-1 إختبار إرتباط الرتب لسبيرمان

The Spearman Rank-Correlation Test

هذا هو أبسط الإختبارات التي يمكن أن تطبق أما الى العينة الصغيرة أو الكبيرة ويمكن ان نضعه بالخطوات الآتية :

الخطوة الأولى : نقوم بإنحدار (Y) على (X) من خلال تقدير المعادلة الإحتمالية الآتية :

$$\hat{Y}_i = \hat{b}_0 + \hat{b}_1 X_i + u_i$$

ونحصل على البواقي (residuals)(ℓ'_s) وهي تقديرات لقيم المتغير العشوائي (u) .

الخطوة الثانية : نرتب البواقي (ℓ)(مع إهمال الإشارة المرافقة لها) ونرتب كذلك قيم المتغير التوضيحي إما تنازلياً أو تصاعدياً وثم نحسب معامل إرتباط الرتب :

$$r_{e.x}{}^{1} = 1 - \frac{6\sum D^2}{n(n^2 - 1)}$$

عندما يكون :

D_i= الفرق بين الرتب المعطاة لكل زوج من (x و e) .

n= عدد المشاهدات في العينة الإحصائية .

إن القيمة العالية لمعامل الرتب (r^1) تقترح وجود مشكلة عدم تجانس التباين ، اذا كان لدينا علاقة مع متغيرات توضيحية عدة فعلينا أن نحسب معامل إرتباط الرتب بين (ℓ_i) وكل واحد من تلك المتغيرات التوضيحية على نحو منفصل .

11-5-2 إختبار كولدفيلد وكوانت (The Goldfeld and Quandt Test)*

إن هذا الإختبار يطبق في حالة العينات الكبيرة والمشاهدات يجب أن يكون عددها على الأقل ضعف عدد المعلمات المطلوب تقديرها في الدالة ، والإختبار يفترض التوزيع الطبيعي (Normality) وعدم الإرتباط المتسلسل لقيم (u) ، والفرضية التي تختبر هذا الإختبار هي :

* **S.M Goldfeld and R.E Quadt ,"Some Tests For Homoscedasticity" Journa of American Statistical Association ,Vo 160 , 1965 , PP.539-47)**

أ- فرضية العدم (Null Hypothesis)

إن قيم تباين المتغير العشوائي متجانسة أو ثابتة (H_0 : U_i's)

ب- الفرضية البديلة (Alternative Hypothesis)

إن قيم تباين المتغير العشوائي غير متجانسة (H_1 : U_i's) (مع تزايد التباين)

أما خطوات هذا الإختبار فهي كما يأتي :

الخطوة الأولى : نرتب المشاهدات إستناداً الى حجم المتغير التوضيحي (X)

الخطوة الثانية : نختار عشوائياً أو إعتباطياً (Arbitrarily) عدداً معيناً (C) من المشاهدات المركزية والتي نحذفها من التحليل ، لقد وجد من بعض التجارب التي قام بها كولدفيلد وكوانت أنه بالنسبة للعينات أكبر من (30) مشاهدة فإن العدد المثالي من المشاهدات المركزية التي يجب أن تحذف من الإختبار هي تقريباً ربع عدد المشاهدات الكلية مثلاً (8) بالنسبة الى (n=30) و (16) بالنسبة الى (n=60) ، والمشاهدات المتبقية (n-c) الى عينتين متساويتين في الحجم ($\frac{n-c}{2}$) أحدهما تتضمن القيم الصغيرة للمتغير التوضيحي (X) والأخرى تتضمن القيم الكبيرة للمتغير التوضيحي (X) .

الخطوة الثالثة : نقدر الإنحدار لكل عينة فرعية ، ونحصل على مجموع تربيع البواقي من واحدة منها :

البواقي من العينة مع القيم المنخفضة لـ(X)($\sum \ell_1^2$)مع درجات حرية ($\left[\frac{n-c}{2}\right]-K$) عندما يكون (K) العدد الكلي للمعلمات في النموذج .

البواقي من العينة الفرعية مع القيم العالية لـ(X) ($\sum \ell_2^2$)مع درجات حرية ($\left[\frac{n-c}{2}\right]-K$) عندما يكون (K) العدد الكلي للمعلمات في النموذج .

واذا قسمنا كل واحد من مجموعات تربيع البواقي أو تربيع الإنحرافات على درجات الحرية الملائمة نحصل على تقـديرات لتباينـات(u'_s) المتغيـر العشوائي في العينتين الإحصائيتين ، وإن نسبة هذين التباينين تكون كما يأتي :

$$F^*=\frac{\dfrac{\sum \ell_2^2}{\left(\dfrac{n-c}{2}\right)-K}}{\dfrac{\sum \ell_1^2}{\left(\dfrac{n-c}{2}\right)-K}}=\frac{\sum \ell_2^2}{\sum \ell_1^2}$$

يكون لها توزيع (F) مع درجات حرية :

$$\upsilon_1=\upsilon_2=\left[\left(\frac{n-c}{2}\right)-K\right]=\left[\frac{(n-c-2K)}{2}\right]$$

عندما يكون :

n= عدد المشاهدات

c= المشاهدات المركزية المحذوفة

K= عدد المعلمات المقدرة من كل إنحدار منهما

فإذا كان التباينين متساويين أو متشابهين ، بمعنـى أن تبـاين المتغيـر العشوائي (u) متجانس (Homoscedastic) أو ثابت ، فإن قيمة (F^*) سـوف تميل الى (1) ، فإذا إختلفت التباينات ، فإن قيمة (F^*) ستكون له قيمـة أكبـر (اذا علمنا أن تصميم الإختبار يتضمن أن ($\sum \ell_2^2$) أكبـر مـن ($\sum \ell_1^2$)) ، نقارن القيمة المشاهدة لـ(F^*) مع القيمة النظريـة (F) مـع درجـات حريـة ($\upsilon_1=\upsilon_2=\left[\frac{(n-c-2K)}{2}\right]$) عند مستوى معنوية مختار ، والقيمـة النظريـة لـ(F) نحصل عليها من الجدول (F) .

فإذا كان :

(F^*) أصغر من (F) الجدولية ، فإننا نقبل أن تباين المتغير العشوائي (u) متجانس ، وكلما كانت قيمة (F^*) المشاهدة (نسبة التباين الثاني الى التباين الأول) أكبر فإن عدم تجانس التباين يكون أقوى في المتغير العشوائي .

11-6 الحلول لمشكلة عدم ثبات التباين

اذا تم إثبات وجود مشكلة عدم تجانس التباين من خلال أو على أساس اي من الإختبارات ، فإن الحل الملائم هو تحويل النموذج الأصلي بطريقة للحصول على شكل نموذج يكون فيه حد الإضطراب المحول أو المتغير العشوائي المحول بتباين ثابت .

وبعد ذلك يمكن أن نطبق طريقة المربعات الصغرى الكلاسيكية على النموذج المحول ، إن النموذج بصيغة التحول يصغر أو يقلل الى التعديلات على البيانات الأصلية ، وهذا التعديل على النموذج يعتمد على الشكل الخاص من تجانس التباين ، بمعنى أن ذلك يعتمد على شكل العلاقة بين تباين المتغير العشوائي (δ_u^2) وقيم المتغير التوضيحي أو المتغيرات التوضيحية .

$$\delta_u^2 = f(X_i)$$

وعموماً فإن تحول النموذج الأصلي (Transformation) يتكون من تقسيم العلاقة الأصلية على الجذر التربيعي للحد أو الصيغة المسؤولة عن تجانس التباين ، ولتوضيح هذه العبارة بأمثلة : إفترض إن النموذج الأصلي هو:

$$\hat{Y}_i = \hat{b}_0 + \hat{b}_1 X_i + u_i$$

يكون فيه (u_i) غير ثابت أو غير متجانس التباين ، ولكنه يفي (Satisfies) ببقية الإفتراضات الإحتمالية لنموذج الإنحدار الخطي .

الحالة الأولى : إفترض إن عدم تجانس التباين هو بصيغة :

$$E(U_i)^2 = \delta_{ui}{}^2 = K^2X^2$$

(عندما يكون (K) ثابت نهائي يتم تقديره مع النموذج) ، بمعنى أن تباين (u) يزداد نسبياً مع (X^2).ونحل من أجل الحصول على قيمة (X^2) يكون لدينا :

$$K^2 = \frac{\delta_{ui}{}^2}{X^2}$$

وهذا يقترح أن التحول الملائم من النموذج الأصلي هو تقسيم العلاقة الأصلية على : ($\sqrt{X^2} = X$)

والذي يعني أن صيغة التحول الملائمة هي :

$$\frac{Y_i}{X_i} = \frac{b_0}{X_i} + b_1 + \frac{u_i}{X_i}$$

إن المتغير العشوائي المحول الجديد ($\frac{u_i}{X_i}$) هو بتباين متجانس طالما أن :

$$E\left(\frac{u_i}{X_i}\right)^2 = \frac{1}{X_i{}^2}E(U_i)^2 = \frac{1}{X_i{}^2}\delta_{ui}{}^2$$

الحالة الثانية :إفترض أن شكل عدم التجانس هو :

$$E(U_i)^2 = \delta_{ui}{}^2 = K^2X$$

إن التحول الملائم للنموذج الأصلي يتألف من تقسيم العلاقة الأصلية على ($\sqrt{X^2}$) :

$$\frac{Y}{\sqrt{X}} = \frac{b_0}{\sqrt{X}} + b_1\frac{X}{\sqrt{X}} + \frac{u}{\sqrt{X}}$$

$$(YX^{-\frac{1}{2}}) = b_0X^{-\frac{1}{2}} + b_1X^{\frac{1}{2}} + \frac{u}{\sqrt{X}}$$

إن المتغير العشوائي أو الحد العشوائي المحول $(\frac{u_i}{\sqrt{X}})$ هــو متجــانس التباين (Homoscedastic) ، مع تباين ثابت = (K^2) .

مثال : متوسط الأجور (Y) وعدد العاملين (X) في (30) شركة وتــم إجراء إنحدار (Y) على (X) للعينة كلها نحصل :

$$\hat{Y} = 7.5 + 0.009X \qquad R^2 = 0.90$$
$$(40.27) \quad (16.10)$$

نتائج أنحدار (Y) على (X) للإثنتي عشرة مشاهدة الأولى :

$$\hat{Y} = 8.1 + 0.006X \qquad R^2 = 0.66$$
$$(39.4) \quad (4.36) \qquad Ess_1 = 0.507$$

نتائج أنحدار (Y) على (X) للإثنتي عشرة مشاهدة الثانية :

$$\hat{Y} = 6.1 + 0.013X \qquad R^2 = 0.60$$
$$(4.16) \quad (3.89) \qquad Ess_2 = 3.9095$$

$$Ess = \sum \ell_i^2$$

وحيث أن :

$$F^* = \frac{Ess_2}{Ess_1} = \frac{3.9095}{0.507}$$
$$= 6.10$$

وهذه القيمة أكبر من (F) النظرية أو الجدولية عنــد درجــات حريــة $(\upsilon_1 = 10)$ و $(\upsilon_2 = 10)$ هي (2.97) ، إذاً نقبل فرضية العــدم ، إن المتغيــر العشوائي غير متجانس التباين .

الفصل الثاني عشر

أسلوب المصفوفات في تحليل الإنحدار

أسلوب المصفوفات في تحليل الإنحدار

12-1 مقدمــــــة

إن جبر المصفوفات يستعمل على نحو متزايد في التحليـــل الرياضـــي الإحصائي . إن أسلوب المصفوفات يعد ضرورة عملية في تحليـــل الإنحـــدار المتعدد طالما أنه يسمح بأنظمة معادلات واسعة وأعداد كبيرة من البيانات .

وفي هذا الفصل فإننا سنقوم أولاً بدراسة مختصرة لمقدمة فـــي جبـــر المصفوفات (Matrix Algebra) ، وبعد ذلك نطبق طرق المصفوفات علـــى نموذج الإنحدار الخطي البسيط ، ولكن لا بد من التأكيد أن جبـــر المصـــفوفات غير ضروري جداً في تحليل الإنحدار البسيط مع متغير توضيحي واحد ، ولكنه يوفر في حالة إستخدامه هنا أساساً جيداً لفهم تطبيق جبـــر المصـــفوفات علـــى حالات الإنحدار المتعدد (Multiple Regression) .

12-2 المصفوفات (Matrices)

12-2-1 تعريف المصفوفة

تعرف المصفوفة بوصفها عبارة عن مستطيل لنظام من العناصر مرتبة بصيغة صفوف (rows) وأعمدة (Columns) .

مثال على المصفوفة :

الصف الأول	6000	23
الصف الثاني	13000	47
الصف الثالث	11000	35
	العمود الأول	العمود الثاني

إن عناصر (Elements) هذه المصفوفة المعينة هي أرقام تعبر عـــن الدخل (العمود الأول) والعمر (العمود الثاني) لثلاثة أشخاص ، إن عناصــر المصفوفة مرتبة بصف (الشخص) وعمود (صــفة الشــخص) ، وهكــذا فــإن العنصر في الصف الأول والعمود الأول (6000) يمثل دخل الشــخص الأول ، والعنصر في الصف الأول وفي العمود الثاني (23) يعبر عن عمــر الشــخص الأول . إن بُعد (Dimension) المصفوفة هو (3×2) وهذا يعني ثلاثة صفوف مع عمودين ، فإذا اردنا أن نعرض الدخل والعمر لألف (1000) شخص بصيغة مصفوفة كماعبرنا عن ذلك فنحن سوف نحتاج الى مصفوفة (1000×2) .

أمثلة أخرى على المصفوفات :

$$\begin{pmatrix} 1 & 0 \\ 2 & 10 \end{pmatrix} \begin{pmatrix} 4 & 7 & 12 & 16 \\ 3 & 15 & 9 & 8 \end{pmatrix}$$

إن هاتين المصفوفتين لهما بُعد (2×2) و (2×4) ، يجب أن نلاحظ أننا دائماً نحدد رقم الصفوف أولاً ثم بعد ذلك نحدد رقم الأعمدة في حالــة إعطــاء أبعاد المصفوفات . وكما هي الحال في الجبر الإعتيادي ربما نســتعمل رمــوز للتعريف بعناصر المصفوفة :

$$\begin{matrix} i=1 \\ i=2 \end{matrix} \begin{pmatrix} J=1 & J=2 & J=3 \\ a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix}$$

لاحظ أن الأرقام المرافقة للحرف (a) يكون الرقم الأول من اليسار يمثل الصف الأول والرقم الثاني من اليسار يمثل العمود ، يجب أن نستعمل الرمــز (a_{ij}) للعنصر في (i^{th}) صف وفي (j^{th}) عمود ، وفي مثالنا أعلاه فإن : ($i=1,2$) و ($j=1,2,3$) .

وربما يشار الى المصفوفة برمز مثل Z , X , A وهكذا فإننا يمكن أن نعرف المصفوفة السابقة كما يأتي :

$$A=\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix}$$

إن الإشارة الى المصفوفة (A) عندئذ يتضمن الإشارة الى نظـــام مـــن (3×2) ، ثمة رمز آخر للمصفوفة (A) السابقة هو :

$$A=[aij] \qquad (j=1,2,3) \text{ و } (i=1,2)$$

وبهذه الطريقة نتجنب الحاجة الى كتابة كل عناصر المصفوفة من خلال كتابة العنصر العام فقط ، إن هذا الترميز يمكن أن يستعمل فقط عنـــدما تكـــون العناصر رموزاً (Symbols) . لنلخص الموضوع فنقول أن مصفوفة مـــع (r) من الصفوف و(c) من الأعمدة سوف تعرض أما بصيغتها الكاملة :

$$A=\begin{pmatrix} a_{11} & a_{12} & a_{ij} & a_{1c} \\ a_{21} & a_{22} & a_{2j} & a_{2c} \\ & & & \\ a_{il} & a_{i2} & a_{ij} & a_{ic} \\ a_{rl} & a_{r2} & a_{rj} & a_{rc} \end{pmatrix}$$

أو بالصيغة المختصرة :

$$A=[aij] \qquad (j=1,\ldots..r) \text{ و } (i=1,2,\ldots.c.)$$

بعض الملحوظات :

1- لا يمكن أن تعتقد أن المصفوفة هي رقم ، إنها منظومة (a set) من العناصر رتبت في نظام ، فقط في حالة المصفوفة ، ذات البعد الواحـــد (1×1) هناك رقم منفرد في المصفوفة .

2- إن الصيغة الآتية لا تمثل مصفوفة :

$$\begin{pmatrix} & 14 & \\ & 8 & \\ 10 & & 15 \\ 9 & & 16 \end{pmatrix}$$

طالما ان الأرقام لم ترتب بصيغة أعمدة وصفوف .

12-2-2 المصفوفة مربعة الأبعاد (Square Matrix)

يقال للمصفوفة مربعة (Square) اذا كـــان عـــدد الصـــفوف (rows) يساوي عدد الأعمدة (clumns) ، وهذين مثالين :

$$\begin{pmatrix} 4 & 7 \\ 3 & 9 \end{pmatrix} \qquad \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

12-2-3 المتجه (Vector)

وهومصفوفة تحتوي على عمود واحد فقـــط وتـــدعى عمـــود متجـــه (Column Cector) أو متجه (Vector) وهذين مثالين :

$$A = \begin{bmatrix} 4 \\ 7 \\ 10 \end{bmatrix} \qquad C = \begin{bmatrix} C_1 \\ C_2 \\ C_3 \\ C_4 \\ C_5 \end{bmatrix}$$

إن المتجه (A) هو مصفوفة (1×3) والمتجه (C) هو مصفوفة (1×5) إن مصفوفة تحتوي على صف واحد فقط تدعى متجه صـــف (row vector) وهنا مثالين أيضاً :

$$B^1 = \begin{bmatrix} 15 & 25 & 50 \end{bmatrix} \qquad D^1 = \begin{bmatrix} C1 & C2 \end{bmatrix}$$

يجب أن نلاحظ أن متجه صف (B^1) (row vector B^1) هو مصفوفة (1×3) ومتجه صف (D^1)(row vector D^1) هو مصفوفة (1×2) .

<u>4-2-12 مقلوب المصفوفة (Transpose)</u>

إن مقلوب المصفوفة (Transpose) لمصفوفة (A) هي مصفوفة أخرى تدعى (A^1) نحصل عليها بوساطة جعل الأعمدة صـــفوفاً والصـــفوف أعمـــدة للمصفوفة (A) .

مثال : اذا كان لدينا المصفوفة (A) :

$$\underset{(2\times3)}{A} = \begin{pmatrix} 2 & 5 \\ 7 & 10 \\ 3 & 4 \end{pmatrix}$$

بعد ذلك فإن المقلوبة (Transpose) لهذه المصفوفة هي :

$$\underset{(2\times3)}{A^1} = \begin{pmatrix} 2 & 7 & 3 \\ 5 & 10 & 4 \end{pmatrix}$$

لاحظ ان العمود الأول للمصفوفة (A) هو الصف الأول للمقلوبة (A^1) وكذلك فإن العمود الثاني للمصفوفة (A) هو الصف الثـــاني للمقلوبـــة (A^1) . بالمقابل فإن الصف الأول من المصفوفة (A) أصبح العمـــود الأول للمقلوبـــة

(A^1) وهكذا ، لاحظ أن أبعاد المصفوفة (A) مؤشر بالأرقام تحت الحرف (A) إن هذه الأبعاد تنعكس في حالة (A^1) .

مثال آخر :

$$\underset{(3\times1)}{C} = \begin{bmatrix} 4 \\ 7 \\ 10 \end{bmatrix} \qquad \underset{(3\times1)}{C^1} = \begin{pmatrix} 4 & 7 & 10 \end{pmatrix}$$

وهكذا فإن المقلوبة (Transpose) لمتجه عمود هي متجه صف ، والعكس بالعكس ، وهذا هو السب الذي جعلنا نستخدم الرمز (B^1) للتعبير عن متجه صف طالما ان من الممكن أن يعبر ذلك عن المقلوبة لمتجه عمود (B) ، وعلى نحوعام لدينا :

$$\underset{r\times c}{A} = \begin{pmatrix} a_{11} & \cdots & a_{1c} \\ \vdots & & \vdots \\ a_{r1} & \cdots & a_{rc} \end{pmatrix} = \left(a_{ij}\right) \begin{matrix} I=1,2,\ldots r \\ J=1,2,\ldots c \end{matrix}$$

$$\underset{r\times c}{A} = \begin{pmatrix} a_{11} & A_{rc} \\ & \\ a_{1c} & a_{rc} \end{pmatrix} = \left(a_{ij}\right) \begin{matrix} j=1,2,\ldots r \\ i=1,2,\ldots c \end{matrix}$$

3-12 تساوي المصفوفات (Equality of Matrices)

إن مصفوفتين يقال أنهما متساويتان اذا كان لهما الأبعاد نفسها وأن كل العناصر المتناظرة متساوية ، اذا تساوت مصفوفتان فإن عناصرهما المتناظرة تكون متساوية .

مثلاً :

$$A = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} \qquad B = \begin{bmatrix} 4 \\ 7 \\ 3 \end{bmatrix}$$

عندئذ فإن تساوي (A) مع (B) أي (A=B) يتضمن أن :

$$a_1 = 4$$

$$a_2 = 7$$

$$a_3 = 3$$

وبالطريقة نفسها اذا كان لدينا المصفوفتين :

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} \qquad B = \begin{bmatrix} 17 & 2 \\ 14 & 5 \\ 13 & 9 \end{bmatrix}$$

عندئذ فإن (A=B) يتضمن :

$$a_{11} = 17 \qquad a_{12} = 2$$

$$a_{21} = 14 \qquad a_{22} = 5$$

$$a_{31} = 13 \qquad a_{32} = 9$$

4-12 أمثلة إنحدار (Regression Examples)

في تحليل الإنحدار فإن واحدة من المصفوفات الأساسية هي متجه (Y) تتألف من (n) من المشاهدات حول المتغير المعتمد (Y) .

$$\underset{n\times 1}{Y} = \begin{bmatrix} Y_1 \\ Y_2 \\ \cdot \\ \cdot \\ Y_n \end{bmatrix}$$

يجب أن نلاحظ أن (Y^1)(مقلوبة) هو متجه صف الآتي :

$$Y^1_{1\times n} = \begin{bmatrix} Y_1 & Y_2 \ldots\ldots\ldots Y_n \end{bmatrix}$$

وهناك مصفوفة أساسية أخرى في تحليل الإنحدار هي مصـــفوفة (X) والتي تعرف في تحليل الإنحدار كما يأتي :

$$\underset{n\times 2}{X} = \begin{pmatrix} 1 & X_1 \\ 1 & X_2 \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ 1 & X_n \end{pmatrix}$$

إن مصفوفة (X) تتألف من عمود الــ(1) وعمود ثاني يحتوي على قيم المتغير التوضيحي (X_i) ، لاحظ أن مقلوبة مصفوفة (X_i) هي :

$$\underset{2\times n}{X^1} = \begin{pmatrix} 1 & 1 & 1 \ldots\ldots\ldots\ldots 1 \\ X_1 & X_2 & X_3 \ldots\ldots\ldots X_n \end{pmatrix}$$

مثال من الإنحدار :

الجدول الآتي يبين أو يعرض بيانات عن المتغير المعتمد (Y) والمتغير التوضيحي (X) :

المشاهدة	X_i	Y_i
1	30	73
2	20	50
3	60	128
4	80	170
5	40	87
6	50	108
7	60	135
8	30	69
9	20	56
10	60	132

إن عرض هذه البيانات بصيغة المصفوفات كما يأتي :

$$Y = \begin{pmatrix} 73 \\ 50 \\ . \\ . \\ . \\ . \\ 132 \end{pmatrix} \qquad X = \begin{pmatrix} 1 & 30 \\ 1 & 20 \\ . & . \\ . & . \\ . & . \\ . & . \\ 1 & 60 \end{pmatrix}$$

12-5 جمع وطرح المصفوفات (Matrix Addition and Subtraction)

إن إضافة أو طرح مصفوفتين يتطلب ان تكونا ذا أبعاد متســاوية ، إن مجموع أو فرق مصفوفتين هو مصفوفة أخرى عناصرها تتكون من مجموع أو فرق العناصر المتناظرة للمصفوفتين .

إفترض أن :

$$\underset{3\times2}{A} = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix} \qquad B = \begin{pmatrix} 1 & 2 \\ 2 & 3 \\ 3 & 4 \end{pmatrix}$$

بعد ذلك :

$$\underset{3\times2}{A + B} \begin{pmatrix} 1+1 & 4+2 \\ 2+2 & 5+3 \\ 3+3 & 6+4 \end{pmatrix} = \begin{pmatrix} 2 & 6 \\ 4 & 8 \\ 6 & 10 \end{pmatrix}$$

وبالطريقة نفسها فإن :

$$\underset{3\times 2}{A - B}\begin{pmatrix} 1-1 & 4-2 \\ 2-2 & 5-3 \\ 3-3 & 6-4 \end{pmatrix} = \begin{pmatrix} 0 & 2 \\ 0 & 2 \\ 0 & 2 \end{pmatrix}$$

وعلى نحو عام :

$$\underset{r\times c}{A=}\left[a_{ij} \right] \quad \underset{r\times c}{B=}\left[b_{ij} \right] \quad \begin{matrix} I=1,2,\ldots r \\ J=1,2,\ldots c \end{matrix}$$

بعد ذلك فإن :

$$\underset{r\times c}{A+B=}\left[a_{ij} + b_{ij} \right] \qquad \underset{r\times c}{A-B=}\left[a_{ij} - b_{ij} \right]$$

مثال من الإنحدار :

إن نموذج الإنحدار الآتي : $i=1,2,\ldots n \qquad Y_i=E(Y_i)+e_i$

يمكن ان يكتب :

$$\begin{bmatrix} Y_1 \\ Y_2 \\ . \\ . \\ Y_n \end{bmatrix} = \begin{bmatrix} E(Y_1) \\ E(Y_2) \\ . \\ . \\ E(Y_n) \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ . \\ . \\ e_n \end{bmatrix}$$

وهكذا فإن مصفوفة (Y) تساوي مجموع مصفوفتين :

مصفوفة تحتوي على القيم المتوقعة للمتغير المعتمد (Y) .

ومصفوفة أخرى تحتوي على قيم الخطأ (Error Terms) أو قيم المتغير العشوائي (ei) .

12–2 ضرب المصفوفات (Matrix Multiplication)

أولاً : ضرب المصفوفة برقم إعتيادي (A Scalar) أو رمز يمثل رقماً

إن الــ(Scalar) هو ليس مصفوفة ، غالباً ما نواجه ضرب مصــفوفة برقم إعتيادي ، وفي هذا المجال فإن كل عنصر في المصفوفة يضرب بــالرقم الإعتيادي الــ(Scalar) ، مثلاً إفترض المصفوفة (A) معطاه كما يأتي :

$$A = \begin{pmatrix} 2 & 7 \\ 9 & 3 \end{pmatrix}$$

وبعد ذلك فإن (4A) عندما يكون (4) هو رقم إعتيادي والنتيجــة كمــا يأتي :

$$4A = \begin{pmatrix} 2 & 7 \\ 9 & 3 \end{pmatrix} = \begin{pmatrix} 8 & 28 \\ 36 & 12 \end{pmatrix}$$

وبالطريقة نفسها فإن ($\mathcal{T}$A) تساوي :

$$\mathcal{T}A = \begin{pmatrix} 2 & 7 \\ 9 & 3 \end{pmatrix} = \begin{pmatrix} 2\mathcal{T} & 7\mathcal{T} \\ 9\mathcal{T} & 3\mathcal{T} \end{pmatrix}$$

عندما تكون ($\mathcal{T}$) رقم إعتيادي أو رمز .

اذا كان هناك عنصراً مشتركاً بين عناصر المصفوفة فإن ذلك العنصر المشترك يمكن أن يكون خارج المصفوفة ويعامــل بوصــفه رقمــاً إعتياديــاً (A Scalar) .

مثلاً :

$$\begin{pmatrix} 9 & 27 \\ 15 & 18 \end{pmatrix} = 3 \begin{pmatrix} 3 & 9 \\ 5 & 6 \end{pmatrix}$$

وبالطريقة نفسها فإن :

$$\begin{bmatrix} \frac{2}{T} & \frac{2}{T} \\ \frac{3}{T} & \frac{8}{T} \end{bmatrix} = \frac{1}{T}\begin{bmatrix} 5 & 2 \\ 3 & 8 \end{bmatrix}$$

وعموماً فإنه اذا كان :

$$A = (a_{ij})$$

وإن (T) رقماً إعتيادياً فإننا نحصل على :

$$T A = T A = (T a_{ij})$$

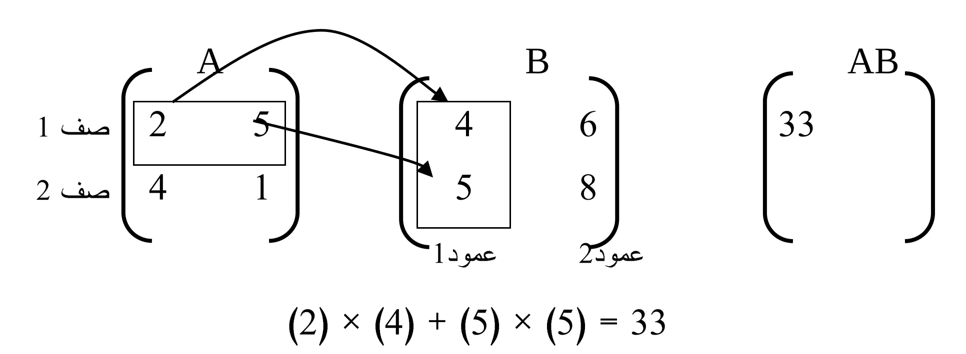
<u>ثانياً : ضرب مصفوفة بمصفوفة أخرى .</u>

إن ضرب مصفوفة بمصفوفة أخرى ربما يبدو معقداً بعض الشيء في البداية ولكن الممارسة ستجعل المسألة روتينية ، لنأخذ المصفوفتين الاتيتين :

$$\underset{2\times2}{A} = \begin{pmatrix} 2 & 5 \\ 4 & 1 \end{pmatrix} \quad \underset{2\times2}{B} = \begin{pmatrix} 4 & 6 \\ 5 & 8 \end{pmatrix}$$

إن حاصل الضرب ينتج (AB) الذي سيكون عبارة عن مصفوفة أبعادها (2×2) وعناصرها تحسب من خلال إيجاد ناتج ضرب صفوف (rows) المصفوفة (A) في أعمدة المصفوفة (B) وجمع تلك النواتج .

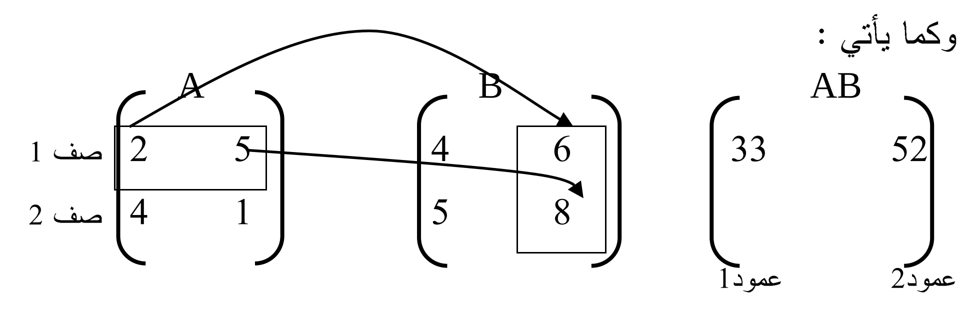
مثلاً لإيجاد عناصر الصف الأول والعمود الأول لحاصل ضرب (AB) فإننا نعمل الصف الأول من (A) والعمود الأول من (B) كما يأتي :



$$(2) \times (4) + (5) \times (5) = 33$$

إن الرقم (33) هو عنصر في الصف الأول والعمود الأول من المصفوفة (AB) ولإيجاد عنصر في الصف الأول والعمود الثاني للمصفوفة (AB) ، فإننا نعمل مع الصف الأول لـ(A) والعمود الثاني لـ(B) .

وكما يأتي :



حاصل الضرب يكون كما يأتي :

$$(2)(6) + (5)(8) = 52$$

والعناصر الأخرى تحسب بالطريقة نفسها :

$$(4)(4) + (1)(5) = 21$$

$$AB = \begin{pmatrix} 33 & 52 \\ 21 & \end{pmatrix}$$

$$(4)(6) + (1)(8) = 32$$

$$AB$$

$$\begin{pmatrix} 33 & 52 \\ 21 & 32 \end{pmatrix}$$

لنأخذ مثالاً آخر :

$$A = \begin{pmatrix} 1 & 3 & 4 \\ 0 & 5 & 8 \end{pmatrix} \qquad B = \begin{bmatrix} 3 \\ 5 \\ 2 \end{bmatrix}$$

$$\underset{2\times1}{AB} = \begin{pmatrix} 1 & 3 & 4 \\ 0 & 5 & 8 \end{pmatrix} \begin{bmatrix} 3 \\ 5 \\ 2 \end{bmatrix} = \begin{bmatrix} 26 \\ 41 \end{bmatrix}$$

$$(1)(3) + (3)(5) + (4)(2) = 26$$

$$(0)(3) + (5)(5) + (8)(2) = 41$$

أمثلة أخرى :

مثال أول :

$$\begin{bmatrix} 4 & 2 \\ 5 & 8 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{pmatrix} 4a_1 & 2a_2 \\ 5a_1 & 8a_2 \end{pmatrix}$$

مثال ثاني :

$$\begin{bmatrix} 2 & 3 & 5 \end{bmatrix} \begin{bmatrix} 2 \\ 5 \\ 3 \end{bmatrix} = \begin{bmatrix} 22+ & 32+ & 52 \end{bmatrix} = \begin{bmatrix} 38 \end{bmatrix}$$

مثال ثالث :

$$\begin{pmatrix} 1 & X_1 \\ 1 & X_2 \\ 1 & X_3 \end{pmatrix} \begin{bmatrix} b_0 \\ b_1 \end{bmatrix} = \begin{pmatrix} b_0 + b_1X_1 \\ b_0 + b_1X_2 \\ b_0 + b_1X_3 \end{pmatrix}$$

12-7 أمثلة في إستخدام المصفوفات في تحليل الإنحدار

لنعرف متجه (Vector) المعلمات (b) كما يأتي :

$$b = \begin{bmatrix} b_0 \\ b_1 \end{bmatrix}$$

عندئذ فإن منتوج (Xb) عندما يكون (X) متغير معرف بالصيغة الآتية في مصفوفة (n ×1) :

$$\underset{n2}{X} = \begin{pmatrix} 1 & X_1 \\ 1 & X_2 \\ 1 & X_n \end{pmatrix}$$

$$\underset{n\times 1}{Xb} = \begin{pmatrix} 1 & X_1 \\ 1 & X_2 \\ . & . \\ . & . \\ 1 & X_n \end{pmatrix} \begin{bmatrix} b_o \\ b_l \end{bmatrix} = \begin{pmatrix} b_0+b_1X_1 \\ b_0+b_1X_1 \\ . \\ . \\ b_0+b_1X_n \end{pmatrix}$$

وطالما ان $E(y_i)=b_0+b_1X_i$ فإننا نرى أن (Xb) هـو متجـه القيمـة المتوقعة $(E(y_i))$ لنموذج الإنحدار الخطي البسيط .

هناك حاصل ضرب آخر نكون بحاجة أليه وهو (Y^1Y) عندما يكـون (Y) هو قيمة مشاهدات المتغير المعتمد معرف كما يلي :

$$\underset{n\times1}{Y} = \begin{pmatrix} Y_1 \\ Y_2 \\ . \\ . \\ Y_n \end{pmatrix}$$

$$Y^1Y=\begin{pmatrix} Y_1 & Y_2.....Y_n \end{pmatrix}\begin{pmatrix} Y_1 \\ Y_2 \\ . \\ . \\ Y_n \end{pmatrix} \begin{pmatrix} = Y^2{}_1+Y^2{}_2.....Y^2{}_n \end{pmatrix}$$

$$=\begin{pmatrix} \sum Y_i^2 \end{pmatrix}$$

يجب ان نلاحظ أن (Y^1Y) هي مصفوفة بأبعاد (1×1).

وهكذا فإننا حصلنا على طريقة لكتابة مجموع تربيع (Y)

$$Y^1Y = \sum Y_i^2$$

كما أننا سوف نحتاج الى (X^1X) هذه مصفوفة بأبعاد (2×2) .

$$\underset{2\times2}{X^1X}=\begin{pmatrix} 1 & 1...........1 \\ X_1 & X_2......X_n \end{pmatrix}\begin{pmatrix} 1 & X_1 \\ 1 & X_2 \\ . & . \\ . & . \\ 1 & X_n \end{pmatrix} = \begin{bmatrix} n & \sum X_i \\ \sum X_i & \sum X_i^2 \end{bmatrix}$$

وكذلك سوف نحتاج الى (X^1Y) وهو مصفوفة بأبعاد 2×1 :

$$\underset{2\times1}{X^1Y}=\begin{pmatrix} 1 & 1\ldots\ldots\ldots..1 \\ X_1 & X_2\ldots\ldots X_n \end{pmatrix}\begin{pmatrix} Y_1 \\ Y_2 \\ . \\ . \\ Y_n \end{pmatrix} = \begin{bmatrix} \sum Y_i \\ \sum X_i Y_i \end{bmatrix}$$

12-8 أنواع خاصة من المصفوفات

12-8-1 المصفوفة المتشابهة (Symmetric Matrix)

اذا كانت $(A^1=A)$ فإن (A) يقال عنها أنها متشابهة وهكذا فـــإن (A) الآتية متشابهة :

$$A=\begin{pmatrix} 1 & 4 & 6 \\ 4 & 2 & 5 \\ 6 & 5 & 3 \end{pmatrix} \quad A`=\begin{pmatrix} 1 & 4 & 6 \\ 4 & 2 & 5 \\ 6 & 5 & 3 \end{pmatrix}$$

إن المصفوفة المتشابهة بالضرورة مصـــفوفة مربعـــة والمصـــفوفات المتشابهة تظهر بصيغة نموذج في تحليل الإنحدار قبل أن نضرب المصـــفوفة (X) في مقلوبتها (X^1)(Transpose) .

والمصفوفة الناتجة (X^1X) .

12-8-2 مصفوفة القطرية ((الداياكنال)) (Diagonal Matrix)

إن مصفوفة القطر الداياكنال (X) هي مصفوفة مربعة تكون عناصرها غير الموجودة على الداياكنال جميعها صفراً ، مثلاً :

$$A=\begin{pmatrix} a11 & 0 & 0 \\ 0 & a22 & 0 \\ 0 & 0 & a33 \end{pmatrix} \qquad B=\begin{pmatrix} 4 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 10 & 0 \\ 0 & 0 & 0 & 5 \end{pmatrix}$$

12-8-3 معكوس المصفوفة (Inverse of Matrix)

في علم الجبر الإعتياددي فإن المعكوس (Inverse) لرقم معين هو الرقم المتبادل معه (Reciprocal) . وهكذا فإن معكوس الرقم (6) هو ($\frac{1}{6}$) .

إن الرقم المضروب في معكوسه يساوي دائماً (1) :

$$6.\frac{1}{6}=1$$

$$X.\frac{1}{X}=X.X^{-1}=1$$

وفي جبر المصفوفات فإن معكوس المصفوفة (A) هي مصفوفة أخرى يشار أليها بــ(A^{-1}) مثل :

$$\mathbf{A^{-1}A=AA^{-1}=I}$$

عندما يكون (I) هو متطابقة المصفوفة . وهكذا فإن متطابقة المصفوفة (I) تلعب الدور نفسه الذي يلعبه الرقم (1) في الجبر الإعتيادي .

إن معكوس المصفوفة (Inverse) يعرف فقط للمصفوفات المربعة . وحتى بالنسبة للمصفوفات المربعة فإن عدة مصفوفات مربعة ليس لها معكوس :

أمثلة : 1– معكوس المصفوفة :

$$\mathbf{A}=\begin{bmatrix} \mathbf{2} & \mathbf{4} \\ \mathbf{3} & \mathbf{1} \end{bmatrix}$$

هو :

$$A^{-1}=\begin{pmatrix} -.1 & 4. \\ .3 & -.2 \end{pmatrix}$$

طالما أن :

$$A^{-1}A=\begin{pmatrix} -.1 & 4. \\ .3 & -.2 \end{pmatrix}\begin{pmatrix} 2 & 4 \\ 3 & 1 \end{pmatrix}=\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

أو :

$$AA^{-1}=\begin{pmatrix} 2 & 4 \\ 3 & 1 \end{pmatrix}\begin{pmatrix} -.1 & .4 \\ .3 & -.2 \end{pmatrix}=\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

2– معكوس المصفوفة

$$A=\begin{pmatrix} 3 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

هو :

$$A^{-1}=\begin{pmatrix} \frac{1}{3} & 0 & 0 \\ 0 & \frac{1}{4} & 0 \\ 0 & 0 & \frac{1}{2} \end{pmatrix}$$

طالما أن :

$$A^{-1}A=\begin{pmatrix} \frac{1}{3} & 0 & 0 \\ 0 & \frac{1}{4} & 0 \\ 0 & 0 & \frac{1}{2} \end{pmatrix}\begin{pmatrix} 3 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

$$=\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

12-8-4 إيجاد معكوس المصفوفة

إن إيجاد معكوس المصفوفة (The Inverse) يمكن غالباً مـــا يتطلـــب عمليات حسابية كثيرة . وهنا سنكتفي بمعكوس المصفوفة (2×2) والمصـــفوفة (3×3) أي كيفية حسابها . أما المصفوفات الأكبر فيعتمد في حســـاب معكـــوس المصفوفة لها على حاسبات إلكترونية ضخمة .

يمكن أن نبين أن معكوس المصفوفة (2×2) أو(3×3) يحسب كما يأتي اذا كانت :

$$A=\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

عندئذ فإن :

$$A^{-1} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \begin{bmatrix} \frac{d}{D} & \frac{-b}{D} \\ \frac{-c}{D} & \frac{a}{D} \end{bmatrix}$$

عندما يكون :

ad-bc=D

D هو محدد (Deteminant) المصفوفة (A)

مثال : أوجد معكوس المصفوفة الآتية :

$$A = \begin{bmatrix} 2 & 4 \\ 3 & 1 \end{bmatrix}$$

لدينا : a=2 b=4

c=3 d=1

D=ad-bc

=(2)(1)-(4)(3)=-10

ولذلك :

$$A^{-1} = \begin{bmatrix} \frac{1}{-10} & \frac{-4}{-10} \\ \frac{-3}{-10} & \frac{2}{-10} \end{bmatrix} = \begin{bmatrix} -.1 & .4 \\ .3 & -.2 \end{bmatrix}$$

مثال على الإنحدار :

إن معكوس المصفوفة في تحليل الإنحدار هو معكوس (Inverse) المصفوفة (X^1X) السابقة :

$$\underset{2\times 2}{X^1X} = \begin{bmatrix} 1 & 1 & . & . & . & 1 \\ X_1 & X_2 & . & . & . & X_n \end{bmatrix} \begin{bmatrix} 1 & X_1 \\ 1 & X_2 \\ . & . \\ . & . \\ . & . \\ 1 & X_n \end{bmatrix} \begin{bmatrix} n & \sum X_i \\ \sum X_i & \sum X_i^2 \end{bmatrix}$$

وبإستخدام قاعدة حساب المعكوس (Inverse) :

$$a = n \qquad b = \Sigma X_i$$

$$c = \Sigma X_i \qquad d = \Sigma_i^2$$

إذاً :

$$D = n\Sigma X_i^2 - (\Sigma X_i)(\Sigma X_i)$$

$$= n\Sigma X_i^2 - (\Sigma X_i)^2$$

$$= n\left(\Sigma X_i^2 - \frac{(\Sigma X_i)^2}{n}\right)$$

$$= n\Sigma(X_i - \overline{X})^2$$

12-8-5 إيجاد معكوس المصفوفة بثلاث صفوف وثلاث أعمدة .

اذا كان لدينا المصفوفة الآتية :

$$X = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & k \end{bmatrix}$$

فإن معكوس هذه المصفوفة :

$$X^{-1} = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & k \end{bmatrix}^{-1} = \begin{bmatrix} A & B & C \\ D & E & F \\ G & H & K \end{bmatrix}$$

عندما يكون :

$$A = \frac{(ek - fh)}{Z} \quad D = \frac{-(dk - fg)}{Z} \quad G = \frac{(dh - eg)}{Z}$$

$$B = \frac{(bk - ch)}{Z} \quad E = \frac{(ak - cg)}{Z} \quad H = \frac{-(ah - bg)}{Z}$$

$$C = \frac{(bf - ce)}{Z} \quad F = \frac{-(af - cd)}{Z} \quad K = \frac{(ae - bd)}{Z}$$

$$Z = a(ek - fh) - b(dk - fg) + c(dh - eg)$$

12-9 مثال على تحليل الإنحدار المتعدد بمتغيرين توضيحيين بإستخدام المصفوفات .

النموذج الآتي يتضمن جدول البيانات حول المبيعات حسب المناطق (Y_i) وهو المتغير المعتمد ، وبيانات حول المتغير التوضيحي عدد السكان في المنطقة (X_1) والمتغير التوضيحي الثاني الفردي (per capita) في المنطقة (X_2) . وفي هذه الحالة فإن النموذج من الدرجة الأولى هو :

$$\mathbf{Y_i = b_0 + b_1X_{i1} + b_2X_{i2}} + \varepsilon_i$$

جدول البيانات

المنطقة	المبيعات بالعدد Y_i	السكان بآلاف الأفراد X_{i1}	الدخل الفردي X_{i2}
1	162	274	2450
2	120	180	3254
3	223	375	3802
4	131	205	2838
5	67	86	2347
6	169	265	3782
7	81	98	3008
8	192	330	2450
9	116	195	2137
10	55	53	2560
11	252	430	4020
12	232	372	4427
13	144	236	2660
14	103	157	2088
15	212	370	2605

حسابات أساسية :

أولاً :

$$1.\ X^1X = \begin{bmatrix} 1 & 1 & . & . & . & 1 \\ 274 & 180 & . & . & . & 370 \\ 2450 & 2254 & . & . & . & 2605 \end{bmatrix} \begin{bmatrix} 1 & 274 & 2450 \\ 1 & 180 & 3254 \\ . & . & . \\ . & . & . \\ . & . & . \\ 1 & 370 & 2605 \end{bmatrix}$$

والذي ينتج عن هذه العملية هو :

$$X^1X = \begin{bmatrix} 15 & 3626 & 44428 \\ 3626 & 1067614 & 11419181 \\ 44428 & 11419181 & 139063428 \end{bmatrix}$$

ثانياً :

$$2.\ X^1Y = \begin{bmatrix} 1 & 1 & . & . & . & 1 \\ 274 & 180 & . & . & . & 370 \\ 2450 & 3254 & . & . & . & 2605 \end{bmatrix} \begin{bmatrix} 162 \\ 120 \\ . \\ . \\ . \\ 212 \end{bmatrix}$$

والذي ينتج عن هذه العملية هو :

$$X^1Y = \begin{bmatrix} 2259 \\ 647107 \\ 7096619 \end{bmatrix}$$

ثالثاً :

$$3.\ X^1X^{-1} = \begin{bmatrix} 15 & 3626 & 44428 \\ 3626 & 1067614 & 11419181 \\ 44428 & 11419181 & 139063428 \end{bmatrix}$$

عندما يكون :

$$a = 15 \qquad b = 3626 \qquad c = 44428 \qquad d = 3626$$

$$e = 1067614 \qquad f = 11419181 \qquad g = 44428 \qquad h = 11419181$$

$$k = 139063428$$

$$Z = 1449704406\ 0000$$

إذاً فإن محدد هذه المصفوفة هو (Z) :

$$A = 1246348416$$

$$B = 0.0002129664\ 176$$

وهكذا .

وجبرياً نلاحظ أن X^1X لنموذج الدرجة الأولى مع متغيرين توضيحيين :

$$X^1X = \begin{bmatrix} 1 & 1 & . & . & . & 1 \\ X_{11} & X_{21} & . & . & . & X_{n1} \\ X_{12} & X_{22} & . & . & . & X_{n2} \end{bmatrix} \begin{bmatrix} 1 & X_{11} & X_{12} \\ 1 & X_{21} & X_{22} \\ . & . & . \\ . & . & . \\ . & . & . \\ 1 & X_{n1} & X_{n2} \end{bmatrix}$$

$$X^1X = \begin{bmatrix} n & \sum X_{i1} & \sum X_{i2} \\ \sum X_{i1} & \sum X_{i1}^2 & \sum X_{i1}X_{i2} \\ \sum X_{i2} & \sum X_{i2}X_{i1} & \sum X_{i2}^2 \end{bmatrix}$$

وهكذا في مثالنا نجد أن :

$$n = 15$$

$$\sum X_{i1} = 274 + 180 + ... = 3626$$

$$\sum X_{i1}X_{i2} = (274)(2450) + (180)(3254)...... = 11419181$$

والى آخره .

وأيضاً يجب أن نلاحظ أن قيمة (X^1Y) للنمــوذج الســابق يمكــن أن تحسب كما يأتي :

$$X^1Y = \begin{bmatrix} 1 & 1 & . & . & . & 1 \\ X_{i1} & X_{21} & . & . & . & X_{n1} \\ X_{i2} & X_{22} & . & . & . & X_{n2} \end{bmatrix} \begin{bmatrix} Y_1 \\ Y_1 \\ . \\ . \\ . \\ Y_n \end{bmatrix} = \begin{bmatrix} \sum Y_i \\ \sum X_{i1} Y_i \\ \sum X_{i2} Y_i \end{bmatrix}$$

مثلاً يكون لدينا :

$$\sum Y_i = 162 + 120 + ... = 2259$$

$$\sum X_{i1} Y_i = (274)(162) + (180)(120) + ... = 647107$$

$$\sum X_{i2} Y_i = (2450)(162) + (3254)(120) ...... = 7096619$$

ولتقدير معلمات العلاقة نستخدم ما يأتي :

$$b = (X^1X)^{-1} X^1Y$$

$$\begin{bmatrix} b_0 \\ b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} 3.45 \\ 0.496 \\ 0.00920 \end{bmatrix}$$

ولذلك فإن :

$$(X^1X)^{-1} = \begin{bmatrix} \dfrac{\sum X_i^2}{n\sum (X_i - \overline{X})^2} & \dfrac{-\sum X_i}{n\sum (X_i - \overline{X})^2} \\ \dfrac{-\sum X_i}{n\sum (X_i - \overline{X})^2} & \dfrac{n}{n\sum (X_i - \overline{X})^2} \end{bmatrix}$$

ولأن : $\sum X_i = n\overline{X} \leftarrow$

نستطيع أن نبسط أكثر :

$$(X^1X)^{-1} = \begin{bmatrix} \frac{\sum X_i^2}{n\sum(X_i-\overline{X})^2} & \frac{-\overline{X}}{n\sum(X_i-\overline{X})^2} \\ \frac{-\overline{X}}{n\sum(X_i-\overline{X})^2} & \frac{1}{n\sum(X_i-\overline{X})^2} \end{bmatrix}$$

12-9-1 إستخدام معكوس المصفوفة .

إفترض أن لدينا المعادلتين الآتيتين :

$$2x + 4y = 20$$
$$3x + y = 10$$

يمكن كتابة المعادلتين بصيغة المصفوفة كما يأتي :

$$\begin{bmatrix} 2 & 4 \\ 3 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 20 \\ 10 \end{bmatrix}$$

إن حل هاتين المعادلتين هو :

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2 & 4 \\ 3 & 1 \end{bmatrix}^{-1}\begin{bmatrix} 20 \\ 10 \end{bmatrix}$$

$$D = ad - bc$$
$$= 2 - 12 = -10$$

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \frac{1}{-10} & \frac{-4}{-10} \\ \frac{-3}{-10} & \frac{2}{-10} \end{bmatrix}^{-1}\begin{bmatrix} 20 \\ 10 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$$

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -0.1 & 0.4 \\ 0.3 & -0.2 \end{bmatrix}\begin{bmatrix} 20 \\ 10 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$$

ولذلك فإن : 2=X ، 4=y

$$(-0.1)(20) + (0.4)(10) = 2$$
$$(0.3)(20) + (-0.2)(10)\ 4$$

12–10 بعض النظريات المهمة للمصفوفات

النظرية الأولى $A + B = B + A$

النظرية الثانية $(A + B) + C = A + (B + C)$

النظرية الثالثة $(AB)C = A(BC)$

النظرية الرابعة $C(A + B) = CA + CB$

النظرية الخامسة $\mathcal{T}(A + B) = \mathcal{T}A + \mathcal{T}B$

النظرية السادسة $(A^1)^1 = A$

النظرية السابعة $(A + B)^1 = A^1 + B^1$

النظرية الثامنة $(AB)^1 = B^1A^1$

النظرية التاسعة $(AB)^{-1} = B^{-1}A^{-1}$

النظرية العاشرة $(ABC)^{-1} = C^{-1}B^{-1}A^{-1}$

النظرية الحادية عشر $(A^{-})^{-1} = A$

النظرية الثانية عشر $(A1)^{-1}=(A-1)^1$

12–11 المتجهـــــــــــــات والمصـــــــــــــفوفات العشـــــــــــــوائية (Random Vectors and Matrices) .

إن المتجه العشوائي أو المصفوفة العشوائية تحتوي على عناصر تمثـــل متغيرات عشوائية . وهكذا فإن المتجه (Y) هو متجه عشوائي طالمـــا أن (Y_i) هي عناصر متغيرات عشوائية .

توقع المتجه العشوائي أو المصفوفة العشوائية .

لنفترض أن لدينا n=3 مشاهدات وأننا نهتم بمتجه المشاهدات :

$$Y = \begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \end{bmatrix}$$

وهنا نشير الى $E(Y)$ بوصفه متجه :

$$E(Y)=\begin{bmatrix}E(Y_1)\\E(Y_2)\\E(Y_3)\end{bmatrix}$$

مثال في الإنحدار :

لنفترض أن عدد المشاهدات في تطبيقات الإنحدار هي $n=3$ ، وهنا فإن الأخطاء العشوائية الثلاثة (Error Terms) $(\varepsilon_3,\varepsilon_2,\varepsilon_1)$ لكل منها توقع صفر . فإذا عرفنا متجه الخطأ :

$$\varepsilon=\begin{bmatrix}\varepsilon_1\\\varepsilon_2\\\varepsilon_3\end{bmatrix}$$

ويكون لدينا : $E(\varepsilon)=0$

طالما أن :

$$E(\varepsilon)=\begin{bmatrix}E(\varepsilon_1)\\E(\varepsilon_2)\\E\varepsilon_3)\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix}$$

12-12 مصفوفة التباين والتباين المشترك للمتجه العشوائي .

لنأخذ مره أخرى المتجه العشوائي (Y) المؤلف أو المكون من ثلاثة مشاهدات Y_3,Y_2,Y_1 . ولكل متغيرين عشوائيين لهما تباين مشترك (covariance) $\delta(Y_i,Y_j)$. نستطيع أن نجمع هذه في مصفوفة تدعى ((مصفوفة التباين والتباين المشترك للمتجه العشوائي Y)) ويشار له بـ $\delta^2(Y)$

$$\delta^2(Y)=\begin{bmatrix}\delta^2(Y_1)&\delta(Y_1,Y_2)&\delta(Y_1,Y_3)\\\delta(Y_2,Y_1)&\delta^2(Y_2)&\delta(Y_2,Y_3)\\\delta(Y_3,Y_1)&\delta(Y_3,Y_2)&\delta^2(Y_3)\end{bmatrix}$$

يجب أن نلاحظ أن التباينات تقع علـــى الـــدايكنال (main diagonal) بينما التباينات المشتركة $\delta(Y_i, Y_j)$ تقع على الصفوف والأعمدة .

12-13 نموذج الإنحدار الخطي البسيط بصيغ المصفوفات .

نحن الآن جاهزون لتطوير صيغ مصفوفات للإنحدار الخطي البسيط .

يجب أن نتذكر مره أخرى أننا سوف لن نعرض نتائج جديدة ولكن فقط نضع صيغ المصفوفات للنتائج التي تم الحصول عليها في البداية .

يجب أن نبدأ مع نموذج الإنحدار الآتي :

$$Y_i = b_0 + b_1X_i + \varepsilon_i \quad i = 1,2, \ldots\ldots, n$$

وهذا يتضمن ما يأتي :

$$Y_1 = b_0 + b_1X_1 + \varepsilon_1$$
$$Y_2 = b_0 + b_1X_1 + \varepsilon_2$$
$$.$$
$$.$$
$$.$$
$$Y_n = b_0 + b_1X_n + \varepsilon_n$$

لقد عرفنا في البداية أو في الجـــزء الســـابق متجـــه المشـــاهدات (Y) ومصفوفة المتغير التوضيحي (X) ومصـــفوفة المعلمـــات (b) ولنعيـــد هـــذه التعريفات الآن :

$$Y = \begin{bmatrix} Y_1 \\ Y_2 \\ . \\ . \\ . \\ Y_n \end{bmatrix} \quad X = \begin{bmatrix} 1 & X_1 \\ 1 & X_2 \\ . & X_3 \\ . & . \\ . & . \\ 1 & X_n \end{bmatrix} \quad b = \begin{bmatrix} b_0 \\ b_1 \end{bmatrix} \quad \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ . \\ . \\ . \\ \varepsilon_n \end{bmatrix}$$

وهنا نستطيع أن نكتب هذه الصيغة بإختصار :

$$\underset{nx1}{Y} = \underset{nx2}{X}\,\underset{2x1}{b} + \underset{nx1}{\varepsilon_n}$$

وطالما أن :

$$\begin{bmatrix} Y_1 \\ Y_2 \\ . \\ . \\ . \\ Y_n \end{bmatrix} = \begin{bmatrix} 1 & X_1 \\ 1 & X_2 \\ . & X_3 \\ . & . \\ . & . \\ 1 & X_n \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ . \\ . \\ . \\ \varepsilon_n \end{bmatrix} = \begin{bmatrix} b_0 + b_1X_1 + \varepsilon_1 \\ b_0 + b_1X_2 + \varepsilon_2 \\ . \\ . \\ . \\ b_0 + b_1X_n + \varepsilon_n \end{bmatrix}$$

إن عمود الــ(1) في مصفوفة الــ(X) ربما ينظر اليه بوصفه يتكـــون من متغير عشوائي $X_0=1$ في نموذج الإنحدار

عندما يكون : $Y_i = b_0X_0 + b1X_i + \varepsilon_i \quad X_0 = 1$

وهكذا فإن مصفوفة (X) ربما تعتبـــر أنهـــا تحتـــوي متجـــه عمـــود (a Column Vector) لمتغير وهمي (أصم) (a dummy variables) ومتجه عمود آخر يتألف من قيم المتغير التوضيحي (X_i) .

أما بالنسبة للمتغير العشوائي أو حدود الخطأ فإننا نفترض أن :

$E(\varepsilon_i) = 0$ وأن $\delta^2_{(\varepsilon)} = \delta^2$ وأن ε_i هي متغيـــرات عشـــوائية إعتيادية مستقلة . إن الشرط $E(\varepsilon_i) = 0$ بصيغة المصفوفة هو :

$$E\begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ . \\ . \\ \varepsilon_n \end{bmatrix} = \begin{bmatrix} E(\varepsilon_1) \\ E(\varepsilon_2) \\ . \\ . \\ E(\varepsilon_n) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ . \\ . \\ 0 \end{bmatrix}$$

إن الشرط أن المتغير العشوائي له تباين ثابت (δ^2) وإن كل التباينـــات المشتركة $\delta(\varepsilon_i,\varepsilon_j)$ بالنسبة الى $j\neq i$ هي صفر . ولذلك فإن مصـــفوفة التبـــاين والتباين المشترك (Variance-Covariance Matrix) :

$$\delta^2 \underset{n\times n}{(\varepsilon)} = \delta^2 I$$

$$\delta^2(\varepsilon) = \delta^2 \begin{bmatrix} 1 & 0 & 0 & . & . & . & 0 \\ 0 & 1 & 0 & . & . & . & 0 \\ . & . & . & & & & . \\ . & . & . & & & & . \\ . & . & . & & & & . \\ 0 & 0 & 0 & . & . & . & 1 \end{bmatrix} = \begin{bmatrix} \delta^2 & 0 & 0 & 0 & . & . & . & 0 \\ 0 & \delta^2 & 0 & 0 & . & . & . & 0 \\ . & . & . & . & & & & . \\ . & . & . & . & & & & . \\ . & . & . & . & & & & . \\ 0 & 0 & .0 & 0 & . & . & . & \delta^2 \end{bmatrix}$$

وهكذا فإن النوذج بصيغة المصفوفة هو :

$$\mathbf{Y = Xb + \varepsilon}$$

عندما يكون ε متجه قيم المتغير العشوائي الذي يتصف بما يأتي :

$$E(\varepsilon) = 0 \qquad \delta^2_{(\varepsilon)} = \delta^2 I$$

12-14 تقدير معلمات الإنحدار بطريقة المربعات الصغرى .

المعادلات الطبيعية أو الإعتيادية (Normal Equations)

$$nb_0 + b_1 \sum X_i = \sum Y_i$$

$$b_0 \sum X_i + b_1 \sum X_i^2 = \sum X_i Y_i$$

وبصيغة المصفوفة :

$$X^1Xb = X^1Y$$

عندما يكون :

$$b = \begin{bmatrix} b_0 \\ b_1 \end{bmatrix}$$

$$\begin{bmatrix} n & \sum X_i \\ \sum X_i & \sum X_i^2 \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \end{bmatrix} = \begin{bmatrix} \sum Y_i \\ \sum X_i Y_i \end{bmatrix}$$

أو :

$$\begin{bmatrix} nb_0 + b_1\sum X_i \\ b_0\sum X_i + b1\sum X_i^2 \end{bmatrix} = \begin{bmatrix} \sum Y_i \\ \sum X_i Y_i \end{bmatrix}$$

<u>12-14-1 تقدير معلمات الإنحدار .</u>

للحصول على المعلمات المقدرة للإنحدار من المعادلات الطبيعية :

$$\mathbf{X^1Xb = X^1Y}$$

وبطرق المصفوفات نضرب الجانبين من مقلوبة المصفوفة لـــ (X^1X) :

$$\mathbf{(X^1X)^{-1}X^1Xb = (X^1X)^{-1}X^1Y}$$

بحيث نجد :

$$\mathbf{(X^1X)^{-1}X^1X = I}$$
$$\mathbf{b = (X^1X)^{-1}X^1Y}$$

<u>مثال</u> : البيانات الآتية عن متغيرين Y و X في الجدول الآتي :

Y_i	X_i	X_iY_i	X_i^2
73	30	2190	900
50	20	1000	400
128	60	7680	3600
170	80	13600	6400
87	40	3480	1600
108	50	5400	2500
135	60	8100	3600
69	30	2070	900
148	70	……	……
132	60	……	……
1100	500	61800	28400

$$n = 10 \quad \Sigma Y_i = 1100 \quad \Sigma X_i = 500 \quad \Sigma X_i^2 = 28400 \quad \Sigma X_i\, Y_i = 61800$$

$$n\Sigma(X_i - \overline{X})^2 = n\left[\Sigma X_i^2 - \frac{(\Sigma X_i)^2}{n}\right]$$
$$= 10\left[28400 - \frac{(500)^2}{10}\right]$$
$$= 34000$$

ولذلك فإن :

$$(X^1X)^{-1} = \begin{bmatrix} \frac{\Sigma X_i^2}{n\Sigma(X_i - \overline{X})^2} & \frac{-\Sigma X_i}{n\Sigma(X_i - \overline{X})^2} \\ \frac{-\Sigma X_i}{n\Sigma(X_i - \overline{X})^2} & \frac{n}{n\Sigma(X_i - \overline{X})^2} \end{bmatrix}$$
$$= \begin{bmatrix} \frac{28400}{34000} & \frac{-500}{34000} \\ \frac{-500}{34000} & \frac{10}{34000} \end{bmatrix}$$
$$= \begin{bmatrix} 0.83529412 & -0.01470588 \\ -0.01470588 & 0.00029412 \end{bmatrix}$$

وأيضاً :

$$X^1Y = \begin{bmatrix} \Sigma Y_i \\ \Sigma X_i Y_i \end{bmatrix} = \begin{bmatrix} 1100 \\ 61800 \end{bmatrix}$$

$$b = \begin{bmatrix} b_0 \\ b_1 \end{bmatrix} = (X^1X)^{-1}X^1Y = \begin{bmatrix} 0.83529412 & -0.0147 \\ -0.0147 & 0.00029412 \end{bmatrix}\begin{bmatrix} 1100 \\ 61800 \end{bmatrix} = \begin{bmatrix} 10.0 \\ 2.0 \end{bmatrix}$$

15-12 تحليل التباين (Analysis of Variance)

لنجعل متجه (Vector) قيم (Y) المقدرة ($\hat{Y}_i$) يشار اليه بـــ ($\hat{Y}$) :

$$\hat{Y} = \begin{bmatrix} \hat{Y}_1 \\ \hat{Y}_2 \\ . \\ . \\ . \\ \hat{Y}_n \end{bmatrix}$$

وقيمة (Vector) البواقي $e_i = \hat{Y}_i - \hat{Y}_i$ يشار اليه بـــ(e) :

$$e = \begin{bmatrix} e_1 \\ e_2 \\ . \\ . \\ . \\ e_n \end{bmatrix}$$

وبصيغة إشارات المصفوفات فإن :

$$\mathbf{\hat{Y} = Xb}$$

$$\begin{bmatrix} \hat{Y}_1 \\ \hat{Y}_2 \\ . \\ . \\ . \\ \hat{Y}_n \end{bmatrix} = \begin{bmatrix} 1 & X_1 \\ 1 & X_2 \\ . & . \\ . & . \\ . & . \\ 1 & X_n \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \end{bmatrix} = \begin{bmatrix} b_0 + b_1X_1 \\ b_0 + b_1X_2 \\ . \\ . \\ . \\ b_0 + b_1X_n \end{bmatrix}$$

$$\mathbf{e = Y - \hat{Y}}$$

إن مجمـــوع المربعـــات (The Sun of Squares) لتحليـــل التبـــاين بإشارات المصفوفات كما يأتي :

$$SSTO = Y^1Y - n\overline{Y}^2$$

$$SSR = b^1X^1Y - n\overline{Y}^2$$

$$SSE = e^1e = Y^1Y - b^1X^1Y$$

<u>مثال</u> : لإيجاد قيمة (SSE) للمثال السابق بطـــرق المصـــفوفات نحـــن نعرف سابقاً أن :

$$Y^1Y = \Sigma Y_i^2 = 134660$$

ونحن نعرف مسبقاً أن :

$$b = \begin{bmatrix} 10.0 \\ 2.0 \end{bmatrix}$$

$$X^1Y = \begin{bmatrix} 1100 \\ 61800 \end{bmatrix}$$

$$b1X1Y = [10.0 \quad 2.0]\begin{bmatrix} 1100 \\ 61800 \end{bmatrix} = 134600$$

ولذلك فإن :

$$\begin{aligned} SSE &= Y^1Y - b^1X^1Y \\ &= 134660 - 134600 \\ &= 60 \end{aligned}$$

<u>مثال ثاني</u> : لنأخذ البيانات الموجودة في الجدول الآتي الـذي يتضـمن الـدخل والإنفاق الإستهلاكي لدولة ما .

الدخل (X) بملايين الدنانير	الإنفاق الإستهلاكي (Y) بملايين الدنانير
80	**70**
100	**65**
120	**90**
140	**95**
160	**110**
180	**115**
200	**120**
220	**140**
240	**155**
260	**150**

الحل : في حالة وجود متغيرين :

$$\hat{b} = \begin{bmatrix} b_0 \\ b_1 \end{bmatrix}$$

$$X^1X = \begin{bmatrix} 1 & 1 & 1 & . & . & 1 \\ X_1 & X_2 & X_3 & . & . & X_n \end{bmatrix} \begin{bmatrix} 1 & X_1 \\ 1 & X_2 \\ 1 & X_3 \\ . & . \\ . & . \\ 1 & X_n \end{bmatrix} = \begin{bmatrix} n & \sum X_i \\ \sum X_i & \sum X_i^2 \end{bmatrix}$$

ثم أن :

$$X^1Y = \begin{bmatrix} 1 & 1 & 1 & . & . & 1 \\ X_1 & X_2 & X_3 & . & . & X_n \end{bmatrix} \begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \\ . \\ . \\ Y_n \end{bmatrix} = \begin{bmatrix} \Sigma Y_i \\ \Sigma X_i Y_i \end{bmatrix}$$

وبإستخدام البيانات في المثال نحصل على :

$$X^1X = \begin{bmatrix} 10 & 1700 \\ 1700 & 32000 \end{bmatrix}$$

$$X^1Y = \begin{bmatrix} 1110 \\ 205500 \end{bmatrix}$$

وبإستعمال قواعد معكوس المصفوفة فإن معكوس المصفوفة (X^1X) هو :

$$(X^1X)^{-1} = \begin{bmatrix} 0.97576 & -0.005152 \\ -0.005152 & 0.0000303 \end{bmatrix}$$

ولذلك فإن :

$$\hat{b}=\begin{bmatrix}\hat{b}_0\\ \hat{b}_1\end{bmatrix}=\begin{bmatrix}0.97576 & -0.005152\\ -0.005152 & 0.0000303\end{bmatrix}\begin{bmatrix}1110\\ 205500\end{bmatrix}$$

$$\begin{bmatrix}\hat{b}_0\\ \hat{b}_1\end{bmatrix}=\begin{bmatrix}243571\\ 0.5079\end{bmatrix}$$

12-15-1 تقدير قيمة R^2(معامل التحديد) .

من خلال إستخدام القانون الآتي :

$$R^2=\frac{\hat{b}^1X^1Y-N\overline{Y}^2}{Y^1Y-N\overline{Y}^2}$$

$$\hat{b}^1X^1Y=\begin{bmatrix}243571 & 0.5079\end{bmatrix}\begin{bmatrix}1110\\ 205500\end{bmatrix}$$

$$=131409.831$$

$$Y^1Y=132100$$

$$N\overline{Y}^2=123210$$

نضع هذه القيم في القانون السابق لـــ(R^2) فنحصل على ما يأتي :

$$R^2=\frac{\hat{b}^1X^1Y-N\overline{Y}^2}{Y^1Y-N\overline{Y}^2}$$

$$=\frac{131409.831-123210}{132100-123210}$$

$$=0.9224$$

الفصل الثالث عشر

نماذج المعادلات الآنية

نماذج المعادلات الآنية Simultaneous Equations Models

13-1 مقدمـــــة

عندما ننظر الى الدراسات والبحوث القياسية التجريبية في مجال تحليل الظواهر الإقتصادية نجد أن هناك عدداً كبيراً من العلاقـات الإقتصـادية هـي العلاقات المعبر عنها بمعادلة واحدة (The Single-Equation Type) ، لقـد كان هذا هو السبب في تركيزنا على هذا النوع من العلاقات الإقتصـادية فـي الفصول السابقة ، في تلك النماذج ذات المعادلة الواحدة هناك متغير واحد معتمد (Y) مثلاً . ويعرض هذا المتغير المعتمد في دالة خطية تحتوي علـى متغيـر توضيحي واحد أو أكثر مثل (X_1,X_2) (Explanatory Variables) وفي مثل هـذه النمـاذج إفتراضـاً ضـمنياً إن علاقـة السـبب بالتـأثير (Cause-effect relationship) بـين المتغيـر المعتمـد (Y) والمتغيـرات التوضيحية $(X_n \;\ldots\ldots,X_3,X_2,X_1)$ هي علاقة بإتجاه واحد : إذ إن المتغيـرات التوضيحية تمثل السبب (Cause) والمتغير المعتمد هو التأثير (effect) .

ولكن هناك حالات أو مواقف تكون فيهـا العلاقـة بـين المتغيـرات الإقتصـادية ذات إتجـاهين أو طـريقين فـي التـأثير (Tow-way Flow of Influence) وهذا يعني إن متغيراً إقتصـادياً واحـداً يؤثر على متغير إقتصادي واحد أو أكثر وذلك المتغير بدوره يتأثر بالمتغيرات التي أثر عليها . وهكذا في تحليل الإنحدار للعلاقة بين النقود (M) وسعر الفائدة (R) ، فإن طريقة المعادلة الواحدة تفترض ضمنياً ان سعر الفائدة ثابت وتحاول تجد درجة إستجابة النقود المطلوبة للتغيرات الحاصلة في مستويات سعر الفائدة ولكن ماذا يحدث اذا كان سعر الفائدة يعتمد على الطلب على النقود .

في هذه الحالة فإن التحليل الذي يعتمد على المعادلة أو الدالـة الواحـدة غير ملائم وذلك لأن (M) الآن يعتمد على (R) وكذلك (R) يعتمد على (M) ،

وهكذا نحن بحاجة الى الإعتماد على معادلتين (Tow-equations) واحدة تربط (M) الى (R) والأخرى تربط (R) الى (M) . وهذا يقودنا الى إستخدام نماذج المعادلات الآنية والتي تحتوي على أكثر من معادلة إنحدار واحدة ، بحيث يكون هناك معادلـــة واحـــدة لكـــل متغيـــر معتمـــد يـــدخل فـــي علاقـــة متبادلـــة (Interdependent) .

13-2 طبيعة نماذج المعادلات الآنية

في نماذج المعادلة الواحدة كان التركيز على تقدير المعلمات والتنبـــوء بمعدلات قيم المتغير المعتمد (Y) المعتمـــدة علـــى القـــيم الثابتـــة للمتغيـــرات التوضيحية ($X_1,X_2,X_3,\ldots\ldots X_n$) ولكن هذا التحليل بالنسبة لبعض العلاقـــات الإقتصادية لايعني شيئاً ، وذلك لأن تلك العلاقات ربما يعبر عنها علـــى نحـــو أفضل بإتجاهين للسببية ، فالمتغير المعتمد (Y) يتأثر بـــالمتغيرات التوضـــيحية ($X_1,X_2,X_3,\ldots\ldots X_n$) وكذلك المتغيرات التوضيحية تتأثر بـــالمتغير المعتمـــد (Y) ، لذلك من الأفضل أن نجمع مجموعة متغيرات مع بعضها البعض والتـــي يمكن أن تحدد بشكل آني أي في الوقت نفسه أو آنياً مع المجموعة المتبقية مـــن المتغيرات ، وهذا هو ما نقوم به في بناء نماذج المعادلات الآنية ، في مثل هذه النماذج يوجد أكثر من معادلة واحدة (معادلة واحدة لكل مجموعة من المتغيرات المعتمدة أو الداخلية المرتبطة مع بعضها) .

في نماذج المعادلات الآنية لا يمكن تقـــدير المعلمـــات (parameters) لمعادلة واحدة دون أن نأخذ في الحساب المعلومات التـــي تزودنـــا بهـــا بقيـــة المعادلات الأخرى في نظام المعادلات الآنية .

مالذي سيحصل اذا قدرنا المعلمات لكل معادلة بتطبيق طريقة المربعات الصغرى الإعتيادية (OLS) دون النظر الى بقية المعادلات في نظـــام نمـــوذج المعادلات الآنية ؟

من الإفتراضات المهمة جداً في طريقة المربعات الصغرى الإعتيادية (OLS) هو إن المتغيرات التوضيحية أما أن تكون غير إحتمالية (Nonstochastic) أو اذا كانت إحتمالية (stochastic) فهي توزع بمعزل عن المتغير العشوائي أو الخطأ المعياري الإحتمالي .

فإذا لم يحصل أي من هذين الإفتراضين فإن هذا يعني إن تقديرات طريقة المربعات الصغرى (Least-squares estimates) هي ليست فقط متحيزة وإنما أيضاً غير متسقة أو غير مترابطة منطقياً أو أن تقديرات المربعات الصغرى تكون متنافرة وهذا يعني أنه كلما زاد حجم العينة الإحصائية الى ما لا نهاية (indefinitely) ، فإن تقديرات المعلمات لا تميل الى الإقتراب من قيم المجتمع الإحصائي ، وهكذا في النظام الخيالي للمعادلات الآنية الآتي :

$$\mathbf{Y_1 = b_0 + b_1Y_2 + b_2X_1 + U_1}$$
$$\mathbf{Y_2 = b_0 + b_1Y_1 + b_2X_1 + U_2}$$

عندما يكون (Y_1,Y_2) يعتمدان على بعضهما البعض أو المتغيرات الداخلية و (X_1) عبارة عن متغير توضيحي أو متغير خارجي (exogenous Variable) و (U1,U2) المتغير العشوائي أو الإحتمالي ، المتغيران (Y1,Y2) كلاهما إحتماليان (Stochastic) .

لذلك ما لم نبين أن المتغير الإحتمالي التوضيحي (Y_2) يوزع بمعزل عن (U_1) وكذلك نبين أن المتغير الإحتمالي التوضيحي (Y_1) يوزع بمعزل عن (U_2) فإن تطبيق طريقة المربعات الصغرى الإعتيادية (OLS) لهذه المعادلات أعلاه على نحو منفرد ستقودنا الى تقديرات متنافرة وغير منسقة .

13-3 مثال على نماذج المعادلات الآنية

المثال الأول : نموذج العرض والطلب

كما نعرف جيداً أن السعر (P) للسلعة والكمية منها (Q) المباعة يتقرران بتقاطع منحنى الطلب مع منحنى العرض للسلعة ، ومن اجل التبسيط نفترض أن منحنى الطلب ومنحنى العرض هما خطيان (Linear) ونضيف الى معادلات الطلب والعرض المتغير العشوائي (U_1) والمتغير العشوائي (U_2) وهكذا يمكن كتابة دالة الطلب ودالة العرض للسلعة كما يأتي :

دالة الطلب $b_1<0$ $\quad Q^d_t = b_0 + b_1P_t + U_{1t}$

دالة العرض $\alpha_1>0$ $\quad Q^s_t = \alpha_0 + \alpha_1P_t + U_{2t}$

حالة التوازن $\quad Q^d_t = Q^s_t$

عندما يكون :

Q^d = الكمية المطلوبة

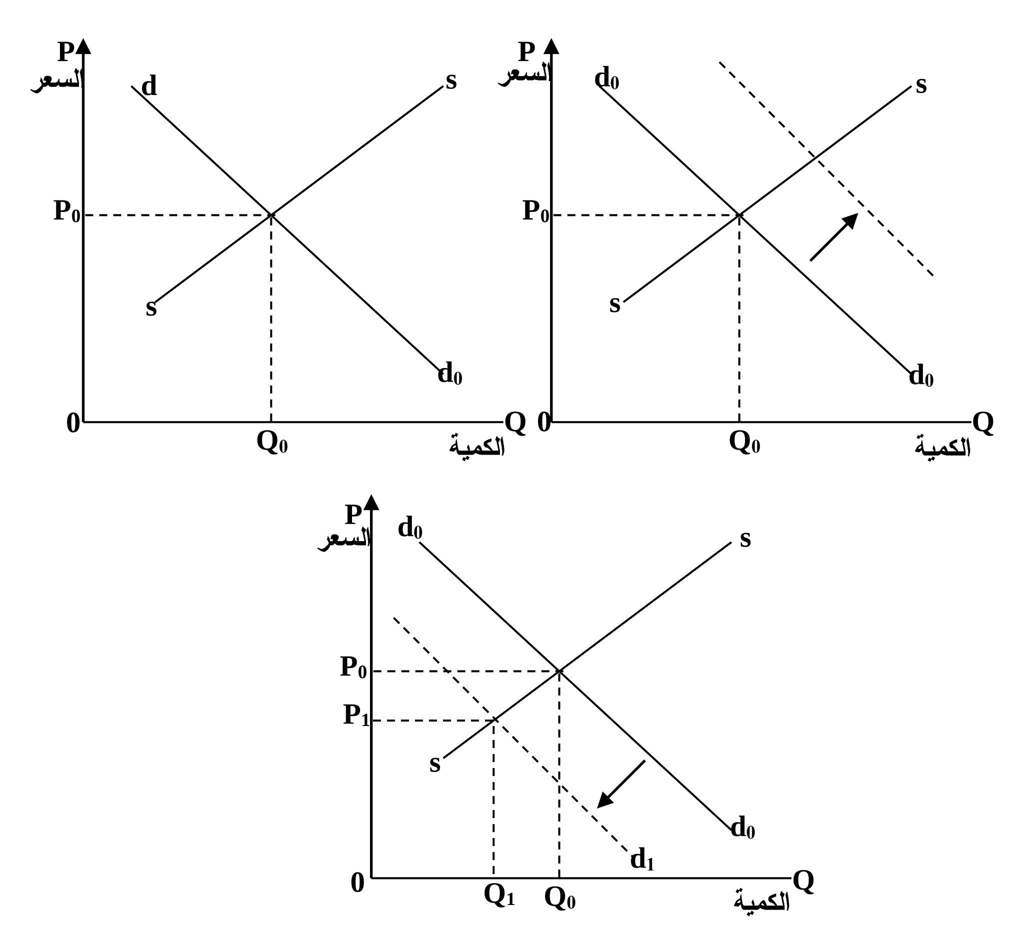
Q^s = الكمية المعروضة

t = الزمن

$\alpha_1, \alpha_0 b_0, b_1$= هي معلمات العلاقة

على معلومات مسبقة فإن (b_1) من المتوقع ان تكون سالبة الإشارة (حيث منحنى الطلب يميل الى الإنحدار الى الأسفل) و (α_1) من المتوقع أن تكون موجبة الإشارة (حيث منحنى العرض يميل الى الإنحدار الى الإرتفاع) .

في هذه الحالة ليس من الصعوبة في ان نرى أن السعر (P) والكمية (Q) يعتمدان على بعضهما البعض ، فإذا تغير (U_{1t}) مثلاً بسبب التغيرات في المتغيرات الأخرى المؤثرة في (Q^d_t) (مثلاً الدخل والثروة والذوق) فإن منحنى الطلب سيتحرك الى الأعلى اذا كان (U_{1t}) موجباً ويتحرك الى الأسفل اذا كان (U_{1t}) سالباً وهذه الحركة مبينة في الشكل (13-1) :



شكل (13-1) الإعتمادية بين السعر والكمية

وكما نرى من الشكل فإن التحرك في منحنى الطلب يغير السعر (P) وكذلك الكمية (Q) ، بالطريقة نفسها فإن التغيير في (U_{1t}) (بسبب الإضطرابات أو المناخ أو القيود على الإستيرادات والصادرات ....الخ) سيحرك منحنى العرض ومرة أخرى هذا التحرك يؤثر في السعر (P) والكمية (Q) ، بسبب هذا الإعتماد الآني (Simultaneous dependence) بين الكمية (Q) والسعر (P) وبين (U_{1t}) و (P_t) وبين (U_{2t}) والسعر (P_t) لا يمكن ان تكون علاقة مستقلة ، لذلك فإن دالة الإنحدار للعلاقة بن (Q) و (P) في معادلة الطلب سوف تخالف

فرضاً مهماً للنموذج الخطي التقليدي (الكلاسيكي) في عدم وجود إرتباط بين المتغيرات التوضيحية والمتغير العشوائي .

المثال الثاني : النموذج الكينزي في تقدير الدخل

لنأخذ نموذجاً كينزياً بسيط في تقرير الدخل :

دالة الإستهلاك $0<b_1<1$ $C_t = b_0 + b_1Y_t + U_t$

متطابقة الدخل $Y_t = C_t + I_t\ (= S_t)$

عندما يكون :

C = الإنفاق على الإستهلاك

Y = الدخل

I = الإستثمار (متغير خارجي)

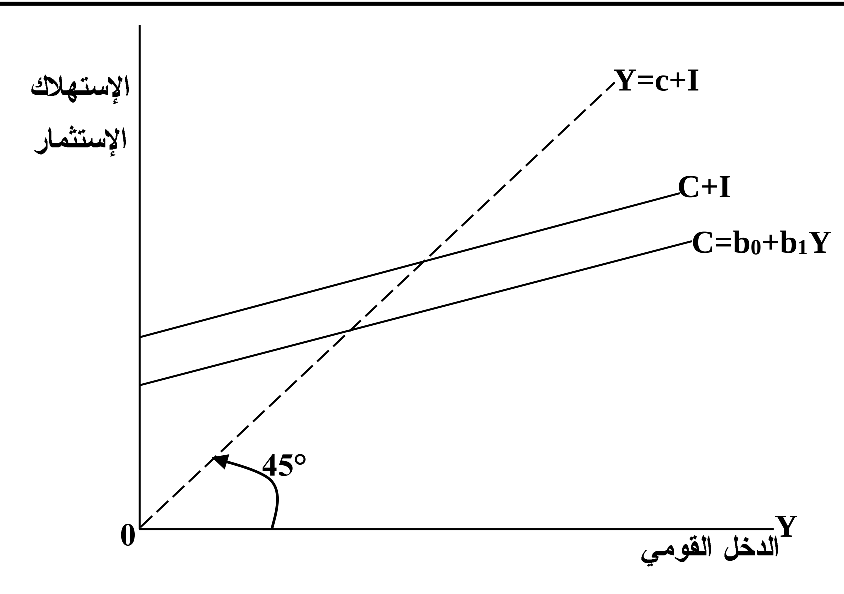
S = الإدخارات

t = الزمن

U = المتغير العشوائي

b_1,b_0 = معلمات دالة الإستهلاك

المعلمة (b_1) تعبر عن الميل الحدي للإستهلاك (MPC) (Marginal Propensity to Consume) وهو الكمية الإضافية من المصروفات الإستهلاكية الناتجة من إضافة دينار واحد الى الدخل . من النظرية الإقتصادية نجد أن (b_1) تقع بين (0) و (1) ، إن دالة الإستهلاك أعلاه هي دالة إحتمالية ومتطابقة الدخل تبين لنا أن الدخل الكلي يساوي المصروف الإستهلاكي الكلي زائداً المصروف الإستثماري الكلي ومن المفهوم في النظرية الإقتصادية أن الإستثمار الكلي يساوي الإدخار الكلي .



شكل (13−2) النموذج الكينزي في تقرير الدخل

من دالة الإستهلاك في الشكل (13−2) واضح ان (C) و (Y) هما متغيران معتمدان على بعض (أي أن هناك درجة من الإعتمادية بينهما) وأن المتغير (Y_t) في دالة الإستهلك من المتوقع أن يكون غير مستقل عن المتغير العشوائي (U_t) وذلك عندما يتحرك (U_t) (بسبب عدد مختلف من العناصر) فإن دالة الإستهلاك أيضاً تتحرك وهذه بدورها تؤثر على (Y_t) .

ولذلك مرة أخرى نقول أن الطريقة التقليدية (الكلاسيكية) للمربعات الصغرى لا يمكن تطبيقها على هذه الدالة ، واذا طبقت فإن التقديرات ستكون متنافرة وغير متطابقة مع المنطق .

13−4 حل تحيز المعادلات الآنية

طالما أن تطبيق طريقة المربعات الصغرى الإعتيادية (OLS) على معادلة تعود الى نظام من المعادلات الآنية تنتج تحيزاً وتقديرات للمعلمات غير متسقة فإننا يجب أن نطبق طرقاً أخرى لتقدير المعلمات في هذه الحالة تعطي تقديرات أفضل للمعلمات ، وهناك طرق عدة لهذا الغرض :

1- طريقة الشكل المختزل أو المصغر

(The Reducd Form)

أو طريقة المربعات الصغرى غير المباشر

2- طريقة المتغيرات الأداتية

(Instrumental Variables)

3- طريقة المربعات الصغرى بمرحلتين

(Two Stage Least Squares 2SLS)

4- طريقة الإحتمال الأعظم بمعلومات محدودة

(Limited Information Maximum Likelihood)

5- طريقة التقديرات المختلطة

(The Mixed Estimation Methods)

6- طريقة المربعات الصغرى بثلاثة مراحل

(Three-Stage LS)

7- الإحتمال الأعظم بمعلومات كاملة

إن الطرق الخمسة الأولى تدعى طرق المعادلة الواحدة (Single-Equation Methods) وذلك لأنها تطبق على معادلة واحدة من النظام في مرة ، أما طريقة المربعات الصغرى بثلاثة مراحل وطريقة الإحتمال الأعظم بالمعلومات الكاملة تدعى طرق الأنظمة (Systems Methods) وذلك لأنها تطبق على كل المعادلات المؤلفة للنظام على نحو آني .

13-4-1 بعض النماذج

1- النماذج التركيبية أو الهيكلية (Structural Models)

إن النموذج التركيبي أو الهيكلي هو نظام كامل أو تام من المعادلات التي تصف هيكل أو تركيب العلاقات بين المتغيرات الإقتصادية ، إن المعادلات

التركيبية أو الهيكلية (Structural Equations) تعبر أو تعرض المتغيـرات الداخلية بوصفها دوال لمتغيرات داخلية أخـرى ولمتغيـرات مقـررة مسـبقاً ومتغيرات عشوائية ، وللتوضيح نستخدم النموذج البسيط لإقتصاد مغلق :

$$C_t = a_0 + a_1 Y_t + U_1$$
$$I_t = b_0 + b_1 Y_t + b_2 Y_{t-1} + U_2$$
$$Y_t = C_t + I_t + G_t$$

إن المعادلة الأولى هي دالة الإستهلاك والمعادلة الثانية هي دالة إستثمار والمعادلة الثالثة هي معادلة تعريفية .

ونظام هذه المعادلات تام وذلك لأنه يحتوي على ثلاثة معـادلات مـع ثلاثة متغيرات داخلية (C_t, I_t, Y_t) ، والنموذج يحتوي على متغيـرين مقـررين مسبقاً وهما المصروف الحكومي (G) والدخل المتخلف زمنيـاً (دخـل السـنة السابقة) (Y_{t-1}) ، ولغرض التبسيط سوف نتجاهل معلمات التقاطع (Intecepts) في المعادلات الهيكلية أو التركيبية .

إن المعلمات التركيبية بشكل عام هي إما ميول أو مرونات أو أي معلمة في النظرية الإقتصادية ، والمعلمة الهيكلية تعرض أو تبـين التـأثير المباشـر للمتغير التوضيحي على المتغير المعتمد . إن الآثار غير المباشـرة يمكـن أن تحسب بحل النظام الهيكلي فقط وليس من خلال المعلمات الهيكلية الفردية ، إن العناصر غير الظاهرة في أية دالة على نحو واضح ربما يكون لها أثـر غيـر مباشرة على المتغير المعتمد لتلك الدالة ، مثلاً التغيير في الإستهلاك سوف يؤثر في الإستثمار على نحو مباشر من خلال تأثير الزيادة في الإستهلاك في الزيادة في الدخل (Y) الذي هو أحد المقررات المهمة للإستثمار .

إن تأثير (C) في (I) لايمكن قياسه مباشرة في أي من المعلمات الهيكلية أو التركيبة ، ولكن يمكن أن نحسب هذا التأثير في الحل الآني للنظام . عادةً أو تقليداً فإن المعلمات الهيكلية يعبر عنها بـ(b'_s) عندما تشير تلك المعلمات الـى

متغيرات داخلية ، ويعبر عنها بــ(γ`s)عندما ترافق هذه المعلمات متغير مقرر مسبقاً ، إن المتغيرات التوضيحية يشار اليها بــالحرف الصــغير (y) ، بينمــا المتغيرات الخارجية يعبر عنها بالحرف الصغير (x'_s) ، وبإستخدام الإشــارات التقليدية فإن النظام الهيكلي (Structural System) يصبح :

$$\mathbf{y_1 = b_{13}y_3 + u_1}$$
$$\mathbf{y_2 = b_{23}y_3 + \gamma`_{21} + u_2}$$
$$\mathbf{y_3 = y_1 + y_2 + x_2}$$

عندما يكون :

$$y_3 = Y$$
$$y_2 = I$$
$$y_1 = C$$
$$x_1 = Y_{t-1}$$
$$x_2 = G$$

وعندما ننقل المتغيرات المشاهدة الى الجهة اليسرى نحصل على جدول كامل للمعلمات الهيكلية أو التركيبية كما يأتي :

$$\mathbf{y_1 + 0y_2 + b_{13}y_3 + ox_1 + 0x_2 = u_1}$$
$$\mathbf{oy_1 + y_2 - b_{23}y_3 - \gamma_{21}x_1 - 0x_2 = u_2}$$
$$\mathbf{-y_1 - y_2 + y_3 + ox_1 - x_2 = 0}$$

2- نماذج الشكل المختزل أو المصغر(Reduced Form Models)

إن هذه النماذج تدعى أيضاً طريقة المربعات الصغرى غير المباشــرة (Indirect Least Squares) ، إن هذه النماذج هي إحــدى طــرق المعادلــة الواحدة (Single-Equation) وفيها يتم إستخدام معادلة واحدة للنظام في كــل مرة ، إن هذه الطريقة ملائمة عندما تكون معادلات النظام الهيكلي أو التركيبي تحتوي على متغيرات مقررة مســبقاً ومتغيــرات توضــيحية بــين مجموعــة المتغيرات ويمكن توضيح طريقة الشكل المختزل بالخطوات الآتية :

الخطوة الأولى : نحصل على الشكل المصغر أو المختزل من النموذج الهيكلي أو التركيبي عن طريق إعادة كتابة المعادلات بطريقة يعبر فيهــا عــن

المتغيرات التوضيحية بوصفها دالة للمتغيرات المقرر مسـبقاً ، وهكـذا فـإن المتغيرات التوضيحية في الشكل المصغر هي متغيرات خارجيـة علـى نحـو حقيقي أو أنها قيم متخلفة زمنياً (Lagged Values) للمتغيرات الداخلية ولذلك فـإن موقفـاً مـن التقريـر المـرتبط أو موقـف السـببية بإتجـاهين (Tow-Way Causation) للمتغيرات يبدو أنه تم تجنبـه بإسـتخدام معادلـة واحدة .

الخطوة الثانية : اذا حصـلنا علـى الإفتراضـات الإعتياديـة حـول الإضطراب (disturbance term) أو المتغير العشوائي لأيـة معادلـة مـن معادلات الشكل المصغر ، عندئذ نطبق طريقة المربعات الصغرى لكل معادلـة بصيغة الشكل المصغر ونحصل على تقديرات لمعادلات الشكل المصغر ، وهذه المعاملات يشار أليها تقليدياً بالحرف اليوناني (الإغريقـي)(π) وهـي خلـيط (a mixture) من معلمات النموذج الهيكلي أو التركيبي .

إن العلاقة بين معاملات الشكل المصغر (π) والمعلمـات التركيبيـة أو الهيكلية (b'_s) ($\gamma`s$) تشكل نظاماً من المعاملات تكون فيه معـاملات الشـكل المصغر معبر عنها بوصفها دوال للمعلمات الهيكلية أو التركيبية .

إن نظام العلاقات بين المعلمات التركيبة (b'_s) ($\gamma`s$) ومعاملات الشكل المختزل (π'_s) ربما تدعى نظام معاملات العلاقات .

الخطوة الثالثة : وهي الخطوة النهائية أو الأخيرة لطريقـة المربعـات الصغرى غير المباشرة تتـألف مـن إسـتخدام تقـديرات معـاملات الشـكل المصغر(π'_s) التي يتم الحصول عليها من الخطوة الثانية أو السابقة وحل نظام معاملات العلاقات للحصول على معلمات تركيبية وإن هذه التقديرات سـتكون وحيدة اذا كان النموذج مشخص تماماً ، ولتوضيح تطبيق طريقـة المربعـات الصغرى غير المباشرة سوف نقوم بأخذ مثال من نظريـة تقريـر أو تحديـد السعر .

نفترض أن آلية السوق لسلعة معينة يمكن وصفها بالنظام الآتـــي مـــن المعادلات الآنية :

$$\mathbf{D = a_0 + a_1P\ a_2Y + U_1}$$
$$\mathbf{S = b_0 + b_1P + b_2W + U_1}$$
$$\mathbf{D = S}$$

عندما يكون :

D = الكمية المطلوبة

S = الكمية المعروضة

P = السعر

Y = الدخل

W = حالة الجو او المناخ

إن هذا النموذج الهيكلي أو التركيبي وهو رياضياً نموذج كامل أو تـــام وذلك لأنه يحتوي ثلاثة معادلات بثلاث متغيرات داخلية (D,S,P) وهذا النظام يحتوي على متغيرين خارجيين هما الدخل (Y) وحالة الجو (W) . إن الشـــكل المصغر أو المختزل للنموذج الذي تعرض فيه المتغيرات الداخلية بوصفها دالة للمتغيرات الخارجية فقط يمكن أن نحصل عليه عن طريق التعويض المستمر .

أنها كما يأتي :

$$\mathbf{b_0 + b_1P + b_2W + U_2 = a_0 + a_1P + a_2Y + u_1}$$
$$\mathbf{b_1P - a_1P = a_0 - a_2Y - b_0 - b_2W + U_1 - U_2}$$
$$\mathbf{P(b_1 - a_1) = a_0 - b_0 + a_2Y - b_2W + U_1 - U_2}$$

عندما يكون $\upsilon_2 = U_1 - U_2$ يكون لدينا المعادلة الآتية :

$$P = \frac{a_0 - b_0}{b_1 - a_1} + \frac{a_2}{b_1 - a_1}Y + \frac{-b_2}{b_1 - a_1}W + \upsilon_2$$
$$D = a_0 + a_1P + a_2Y + U_1$$

$$D = a_0 + a_1\left(\frac{a_0 - b_0}{b_1 - a_1} + \frac{a_2}{b_1 - a_1}Y + \frac{-b_2}{b_1 - a_1}W + \frac{U_1 - U_2}{b_1 - a_1}\right) + a_2Y + U_1$$

$$D = a_0 + \frac{a_1(a_0 - b_0)}{b_1 - a_1} + \frac{a_1a_2}{b_1 - a_1}Y + \frac{-a_1b_2}{b_1 - a_1}W + \frac{a_1(U_1 - U_2)}{b_1 - a_1}$$

$$D = \frac{a_0(b_1 - a_1)}{b_1 - a_1} + \frac{a_1(a_0 - b_0)}{b_1 - a_1} + \frac{a_1a_2}{b_1 - a_1}Y + \frac{-a_1b_2}{b_1 - a_1}W + \frac{a_1(U_1 - U_2)}{b_1 - a_1}$$

$$D = \frac{a_0b_1 - a_0a_1}{b_1 - a_1} + \frac{a_1a_0 - a_1b_0)}{b_1 - a_1} + \frac{a_1a_2}{b_1 - a_1}Y + \frac{-a_1b_2}{b_1 - a_1}W + \frac{a_1(U_1 - U_2)}{b_1 - a_1}$$

$$D = \frac{a_0b_1 - \cancel{a_0a_1} + \cancel{a_1a_0} - a_1b_0}{b_1 - a_1} + \frac{a_1a_2}{b_1 - a_1}Y + \frac{-a_1b_2}{b_1 - a_1}W + \frac{a_1(U_1 - U_2)}{b_1 - a_1}$$

$$D = \frac{a_0b_1 - a_1b_0}{b_1 - a_1} + \frac{a_1a_2}{b_1 - a_1}Y + \frac{-a_1b_2}{b_1 - a_1}W + \frac{a_1U_1 - a_1U_2}{b_1 - a_1} + a_2Y + U_1$$

$$D = \frac{a_0b_1 - a_1b_0}{b_1 - a_1} + \frac{a_1a_2}{b_1 - a_1}Y + a_2Y + \frac{-a_1b_2}{b_1 - a_1}W + \frac{a_1U_1 - a_1U_2}{b_1 - a_1} + U_1$$

$$\frac{a_1a_2}{b_1 - a_1}Y + \frac{a_2Y}{1}$$

$$\frac{a_1a_2Y + (b_1 - a_1)a_2Y}{b_1 - a_1}$$

$$\frac{\cancel{a_1a_2Y} + a_2b_1Y - \cancel{a_1a_2Y}}{b_1 - a_1}$$

$$\frac{a_2b_1}{b_1 - a_1}Y$$

$$= \frac{a_1U_1 - a_1U_2 + u_1(b_1 - a_1)}{b_1 - a_1}$$

$$= \frac{a_1U_1 - a_1U_2 + U_1b_1 - a_1U_1}{b_1 - a_1}$$

$$= \frac{U_1b_1 - a_1U_1}{b_1 - a_1}$$

$$D = \frac{a_0b_1 - a_1b_0}{b_1 - a_1} + \frac{a_2b_1}{b_1 - a_1}Y + \frac{-a_1b_2}{b_1 - a_1}W + \frac{U_1b_1 - a_1U_2}{b_1 - a_1}$$

وبإستخدام الترميز التقليدي لمعلمات الشكل المختـــزل (π's) نحصـــل على :

$$\mathbf{D = \pi_{10} + \pi_{11}Y + \pi_{12}W + \upsilon_1}$$
$$\mathbf{P = \pi_{20} + \pi_{21}Y + \pi_{22}W + \upsilon_2}$$

عندما يمكن أن نبين أن :

$$\pi_{10} = \frac{a_0 b_1 - a_1 b_0}{b_1 - a_1} \qquad \pi_{20} = \frac{a_0 - b_0}{b_1 - a_1}$$
$$\pi_{11} = \frac{a_2 b_1}{b_1 - a_1} \qquad \pi_{21} = \frac{a_2}{b_1 - a_1}$$
$$\pi_{12} = \frac{-a_1 b_2}{b_1 - a_1} \qquad \pi_{22} = \frac{-b_2}{b_1 - a_1}$$

بإستخدام بيانات عينة إحصائية للمتغيرات (W,Y,P,D) ربما نطبق طريقة المربعات الصغرى الإعتيادية (OLS) على معادلات الشكل المختزل (بأخذ معادلة واحدة كل مرة) للحصول على تقديرات لمعلمات الشكل المختزل (π'_s) .

ثم نقوم بإحلال القيم المقدرة لمعلمات الشكل المختزل (π'_s) في نظام علاقات المعاملات فنحصل على المعلمات الهيكلية الآتية :

$$a_0 = \pi_{20}\left(\frac{\pi_{10}}{\pi_{20}} - \frac{\pi_{12}}{\pi_{22}}\right) \qquad a_1 = \frac{\pi_{12}}{\pi_{22}}$$
$$a_2 = \pi_{21}\left(\frac{\pi_{11}}{\pi_{21}} - \frac{\pi_{12}}{\pi_{22}}\right)$$
$$b_0 = \pi_{20}\left(\frac{\pi_{10}}{\pi_{20}} - \frac{\pi_{11}}{\pi_{21}}\right) \qquad b_1 = \frac{\pi_{11}}{\pi_{21}}$$
$$b_2 = \pi_{22}\left(\frac{\pi_{12}}{\pi_{22}} - \frac{\pi_{11}}{\pi_{21}}\right)$$

طالما أن المعادلات الهيكلية مشخصة تماماً ، فإن المعادلات تصف العلاقات بين المعلمات الهيكلية (b'_s) ومعلمات الشكل المختزل (π'_s) وهذه العلاقات تعطينا قيم وحيدة للمعلمات الهيكلية (b'_s) . وفي مثالنا فإن نظام علاقات المعلمات يحتوي على ستة معادلات في ستة مجاهيل (Unknouns) وهي المعلمات الهيكلية (b_2,b_1,b_0)(a_2,a_1,a_0) وبحل النظام يمكن الحصول على معلمات هيكلية مقدرة .

3- نمـــــــــاذج المربعـــــــــات الصـــــــــغرى بمـــــــــرحلتين (2SLS) (Tow-Stage Least Squares)

ثمة أسلوب عام ومباشر للحصول على تقديرات لمعلمات معادلة هيكلية مشخصة . وعلينا أن نستمر بالإفتراض أن لدينا عينة عشوائية من المشـــاهدات على كل من مشاهدات (Y) ومشاهدات (X) . وهذا الأســـلوب هـــو أســـلوب المربعات الصغرى بمرحلتين التي تعطي اسم (2SLS) ، في المرحلة الأولـــى نقدر كل معادلة شكل مختزل بطريقة المربعات الصغرى بالإنحـــدار الخطـــي ونحسب القيم الناتجة من هذا التقدير ، في المرحلة الثانية نأخذ كل معادلة هيكلية (Structural Equation) حسب دورها ونعوض عن كل متغير داخلي يقع في الجانب الأيمن من المعادلة بالقيمة المقدرة له في المرحلة الأولى وبعد ذلك نقوم بتقدير المربعات الصغرى بالإنحدار الخطي .

ولنكون أكثر تحديداً لنعمل عبر نموذج (A) الذي يتضمن الشكل الهيكلي الآتي :

الطلب : $Y_1 = \alpha_1 Y_2 + \alpha_2 + \alpha_3 X_1 + U_1$

العرض : $Y_2 = \alpha_4 Y_1 + \alpha_5 + \alpha_6 X_2 + \alpha_7 X_3 + U_2$

والشكل المختزل لهذا النموذج هو :

الكمية : $Y_1 = \pi_1 + \pi_2 X_1 + \pi_3 X_2 + \pi_4 X_3 + V_1$

السعر : $Y_2 = \pi_5 + \pi_6 X_1 + \pi_7 X_2 + \pi_8 X_3 + V_2$

المرحلة الأولى : تقدير الشكل المختزل نحصل على :

$$\hat{Y}_1 = P_1 + P_2 X_1 + P_3 X_2 + P_4 X_3$$

$$\hat{Y}_2 = P_5 + P_6 X_1 + P_7 X_2 + P_8 X_3$$

ونخلق المتغيرات ($\hat{Y}_1$ و $\hat{Y}_2$) عند كل مشاهدة ($\hat{Y}_1$ و $\hat{Y}_2$)

المرحلة الثانية : بالنسبة لمعادلة الطلب نجري إنحدار (Y_1) على ثلاثة متغيرات $(\hat{Y}_2)$و(1) {وهو المتغير الذي يكون معامله هو معلمة التقاطع} و(X_2) فنحصل على :

$$\tilde{Y}_1 = a_1^* Y_2 + a_2^* + a_3^* X_1$$

وبالنسبة لمعادلة العرض : نجري إنحدار (Y_2) على أربعة متغيرات هي :

$(\hat{Y}_1$ و 1 و X_2 و $X_3)$ فنحصل على :

$$\tilde{Y}_2 = a_4^* Y_1 + a_5^* + a_6^* X_2 + a_7^* X_3$$

إن هذه المعلمات (a_s^*) هي تقديرات طريقة المربعات الصغرى بمرحلتين للمعلمات (α_s) .

فإذا كانت المعادلة الهيكلية (A Structural Equation) مشخصة تماماً فإن تقديرات طريقة المربعات الصغرى بمرحلتين تكون متطابقة أو متسقة مع تقديرات المربعات الصغرى غير المباشرة ، بينما اذا كانت المعادلة الهيكلية فوق التشخيص عندئذ فإن طريقة المربعات الصغرى بمرحلتين تربط بفعالية التقديرات البديلة لطريقة المربعات الصغرى غير المباشرة ، ولكن في كلا الحالتين فإن تقديرات طريقة المربعات الصغرى بمرحلتين (2SLS) تكون متسقة (Consistent) .

<u>مثال : طريقة (2SLS)</u>

لنأخذ نموذج إقتصادي جزئي وهو عن العرض والطلب على العمل :

العرض : $Y_1 = \alpha_1 Y_2 + \alpha_2 + \alpha_3 X_1 + U_1$

الطلب : $Y_2 = \alpha_4 Y_1 + \alpha_5 + \alpha_6 X_2 + \alpha_7 X_3 + \alpha_8 X_4 + U_2$

عندما يكون :

Y_1 = الأشهر المشتغلة (الكمية)

Y_2 = معدل الأجور (السعر)

وهذه هي متغيرات داخلية (Endogenous Variables) أما المتغيرات الخارجية (Exogenous Variables) فضلاً عن متغيرات الواحد (1) فهي :

X1 حجم العائلة X2 التعليم X3 العمر X4 العنصر (الأصل)

وهي متغيرات غيرمستقلة عن حدود الإضطراب الهيكلية (U_2,U_1) إن حدود الإضطراب هذه لها قيم توقعات صفر وربما تكون مرتبطة ببعضها .

إن النظرية الإقتصادية للعرض والطلب تقول أن العمال يرغبون بالعمل اكثر عند اجور أعلى ، بينما الشركات ترغب ان تدفع أجوراً أقل للعمل الإضافي وهكذا فإن (α_1) موجبة $(\alpha_1>0)$ و$(\alpha 4)$ سالبة $(\alpha 4<0)$ ، إن الشكل المختزل لهذا النموذج هو :

$$Y_1 = \pi_1 + \pi_2 X_1 + \pi_3 X_2 + \pi_4 X_3 + \pi_5 X_4 + V_1 \quad \text{: الكمية}$$

$$Y_2 = \pi_6 + \pi_7 X_1 + \pi_8 X_2 + \pi_9 X_3 + \pi_{10} X_4 + V_2 \quad \text{: السعر}$$

عندما يكون :

$$\pi_1 - \alpha_1 \pi_6 = \alpha_2$$

$$\pi_2 - \alpha_1 \pi_7 = \alpha_3$$

$$\pi_3 - \alpha_1 \pi_8 = 0$$

$$\pi_4 - \alpha_1 \pi_9 = 0$$

$$\pi_5 - \alpha_1 \pi_{10} = 0$$

جانب الطلب

$$\pi_6-\alpha_4\pi_1 = \alpha_5$$
$$\pi_7-\alpha_4\pi_2 = 0$$
$$\pi_8-\alpha_4\pi_3 = \alpha_6$$
$$\pi_9-\alpha_4\pi_4 = \alpha_7$$
$$\pi_{10}-\alpha_4\pi_5 = \alpha_8$$

عندما نقوم بتقدير النموذج بطريقة إنحدار المربعات الصغرى المباشرة على المعادلات الهيكلية نحصل على النتائج الآتية :

$$\tilde{\tilde{Y}}_1 = 0.271Y_2 + 11.748 - 0.053X_1$$
$$(0.260) \quad (0.378) \quad (0.073)$$

$$\tilde{\tilde{Y}}_2 = 0.003Y_1 - 0.539 \qquad + 0.067X_2 + 0.011X_3 - 0.119X4$$
$$(0.034) \quad (0.469) \qquad (0.011) \quad (0.005) \quad (0.146)$$

وطبعاً فإن هذه النتائج ينبغي عدم تفسيرها بوصفها تقديرات لمعـــادلات العرض والطلب الهيكلية .

ثم نستمر بالإنتقال الى تقديرات طريقة المربعات الصغرى بمرحلتين ، ففي المرحلة الأولى تقدر معادلات الشكل المختزل السابقة بطريقـــة الإنحـــدار الخطي للمربعات الصغرى فنحصل على :

$$\hat{Y}_1 = 11.930 - 0.064X_1 + 0.041X_2 - 0.010X_3 - 0.628X_4$$
$$(0.912) \quad (0.074) \quad (0.033) \quad (0.015) \quad (0.436)$$

$$\hat{Y}_2 = 0.593 + 0.014X_1 + 0.067X_2 + 0.011X_3 - 0.119X_4$$
$$(0.302) \quad (0.025) \quad (0.011) \quad (0.005) \quad (0.144)$$

وبعد ذلك في المرحلة الثانية نضع هذه القيم المقدرة ($\hat{Y}_1 \quad \hat{Y}_2$) علـــى الجانب اليمن من معادلة العرض ومعادلة الطلب وكما يأتي :

العرض
$$\tilde{Y} = 0.767\hat{Y}_2 + 11.420 - 0.054X_1$$
$$(0.472) \quad (0.464) \quad (0.075)$$

الطلب $\tilde{Y} = -0.218\hat{Y}_1 + 2.013 + 0.076X_2 + 0.009X_3 - 0.259X_4$

$(0.462) \quad (5.330) \quad (0.024) \quad (0.007) \quad (0.334)$

من وجهة النظر الإقتصادية فإن الصفة الأكثر أهمية لتقديرات (2SLS) بالمقارنة مع نتائج طريقة المربعات الصغرى المباشرة هي أن معامل الكمية في معادلة الطلب الآن أصبح لها الإشارة الصحيحة ومعامل السعر في معادلة العرض الآن أكبر .

الفصل الرابع عشر

التشخيص الإحصائي

Statistical Identification

التشخيص الإحصائي (The Identification)

14-1 مقدمـــــــة

من المعروف لدى المختصين بالإقتصاد القياسي أن معرفـــــة الشـــــكل المختزل لنظام من المعادلات ليس كافياً دائماً (ليس شرطاً كافياً) ليسمح لنا أن نتعرف على قيمة المعلمات في المنظومة او المجموعة الأولى من المعـــادلات الهيكلية (Structural Equation) . إن المشكلة المتعلقة في قدرتنا من عدمها على تقرير أو تحديد المعادلات الهيكلية عنـــدما نعـــرف الشـــكل المختـــزل او المختصـــــــر أو نحصـــــــل عليـــــــه تـــــــدعى مشـــــــكلة التشـــــــخيص (Identification Problem) ، ولكن على كل حال يجب أن نلاحظ أن معرفة المعلمات الهيكلية (Parameters Structural) ، ليس ضرورة مطلقة اذا كان التنبوء هو غرضنا الأول ، وذلك لأن التنبوءات يمكن ان نحصل عليهـــا مـــن خلال معادلات الشكل المختزل مباشرة .

إن أهمية مشكلة التشخيص تأتي قبل إعتبار مشكلة التقدير ففي اللحظـــة التي يحدد فيها النموذج الهيكلي ، فإننا يجب أن ندقق فوراً أو حالاً مـــن أجـــل معرفة في ما اذا كانت المعادلة أو المعادلات في النموذج مشخصة ، بمعنى في ما اذا كان بإمكاننا أن نحصل على معرفة المعلمات في اللحظة التي يكون فيها الشكل المختزل قد تم تقديره ، بعد التعامل مع مشكلة التشخيص والإنتهاء منهـــا نستطيع أن نتقدم للإختيار من بين أساليب التقدير .

إن الدالة دون أو أقل من التشخيص اذا لم يكن هناك طريقاً لتقدير جميع المعلمات الهيكلية من الشكل المختزل ، والمعادلة تكون مشخصة اذا كان مـــن الممكن الحصول على قيم المعلمات من معادلة الشكل المختزل ، والمعادلة تكون مشخصة تماماً اذا كانت قيمة وحيدة (Unique Value) للمعلمـــة موجـــودة ،

وتكون المعادلة فوق المشخصة (Overidentified) اذا كان بالإمكان الحصول على أكثر من قيمة واحدة لبعض المعلمات .

لننتقل الى المستوى العملي من خلال أخذ نماذج العرض والطلب ، فإذا أخذنا أولاً نموذج العرض والطلب لسلسلة زمنية ، والذي ليس فيـه متغيـرات مقررة مسبقاً :

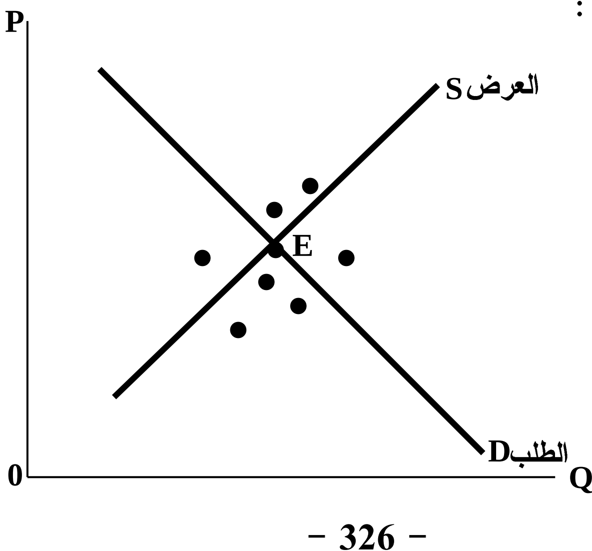
العرض : $_2P_t\alpha_1 + \alpha Q_t = + \Sigma_t$

الطلب : $Q_t = B_1 + B_2P_t + U_t$

نفترض أن السوق في وضع توازني في كل فترة زمنية بحيث أن الكمية المطلوبة تساوي الكمية المعروضة ، إن العنصر الأساسي فـي فهـم مشـكلة التشخيص في سياق هذا النموذج هو ان يتم التركيز على حالة أو شرط التوازن.

في كل فترة زمنية هناك فقط قيمة واحدة للسعر (P) ، وقيمـة واحـدة للكمية المباعة (Q) . بعبارة أخرى إن البيانات المتوافرة للمختص بالإقتصـاد القياسي هي قيم السوق فقط (لكل فترة زمنية) للسعر (P) والكمية (Q) .

إن الأخطاء في المعادلات ستجعل من الممكن لقيم (P) و(Q) المتحصلة غير متشابهة ، ولكن من الممكن أن كل القيم سوف تقـع بـالقرب مـن القـيم التوازنية لـ(P) و(Q) تقرر بالحل المباشر لمعادلات النموذج ، وهذا الموقـف يبدو في الشكل الآتي :



فعندما نحاول أن نقدر معادلة العرض المنفصلة (Separate) ومعادلـــة الطلب المنفصلة بإستخدام بيانات السوق نحصل على نتائج لا معنى لها ، لـــيس هناك إمكانية للتأكد من الميل الحقيقي للطلب والميل الحقيقي للعـــرض عنـــدما تكون بيانات التوازن معروفة فقط . وفي الحقيقة فإن السبب الوحيد الذي يجعل التقدير ممكناً هو أن الأخطاء تظهر في كلا المعادلتين .

إن النموذج الذي نصفه هو نموذج يتضمن أن منحنى العرض ومنحنى الطلب غير مشخصين أو غير معرفين ، وإن كلا المعادلتين غير مشخصتين ، وذلك لأنه ليس هناك طريقاً للحصول على قيم للمعلمات الهيكلية (الميل ومعلمة التقاطع لمنحيي الطلب الفردي ومنحنى العرض الفردي) من معـــادلات الشـــكل المختزل أو المصغر ، إن معادلتي الشكل المختزل :

$$P_t = \frac{U_t - \varepsilon_t}{\alpha_2 - \beta_2} \quad , \quad q_t = \frac{\alpha_2 U_t - \beta_2 \varepsilon_t}{\alpha_2 - \beta_2}$$

يجب أن يكون واضحاً من تفحص الشـــكل الســـابق أن أي زوج مـــن منحنيي العرض والطلب يتقاطعان عند نقطة (E) ومن السهولة أن يبدوان أنهما منحنيي العرض والطلب الحقيقي . بعبارة أخرى هناك عدد غير محـــدود مـــن النماذج الهيكلية (منحنى العرض والطلب) التي تتنسق مـــع الشـــكل المختـــزل نفسه(القيم التوازنية لــ(P) و(Q) .

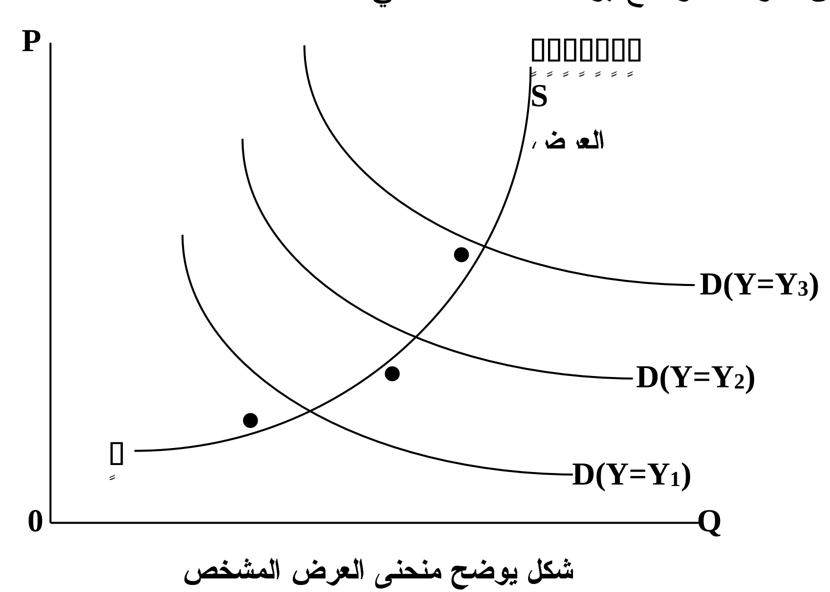
يجب أن يكون واضحاً إن تشخيص معادلات في نظام نموذج يتطلـــب بالضرورة معلومات أكثر من هذا النوع أو ذاك ، لنأخذ مثلاً نظـــام العـــرض والطلب السابق :

العرض : $_2P_t\alpha_1 + \alpha Q_t = + \varepsilon_t$

الطلب : $Q_t = \beta_1 + \beta_2 P_t + \beta_3 Y_t + u_t$

في ظل إفتراض أن $(0 \neq \beta_3)$ وأن (Y_t) تتغير عبر الـــزمن ، فإننـــا لا نستطيع أن نرسم منحنى طلب واحد ومنحنى عرض واحد لكل الفترات الزمنية.

ولإن الدخل يقرر الطلب والدخل يتغير عبر الزمن فإننا يجب أن نأخذ بنظر الإعتبار أو في الحساب الحقيقة التي تقول أن منحنى الطلب ينتقل عبر الزمن ، وهذا الوضع يوضحه الشكل الآتي :



شكل يوضح منحنى العرض المشخص

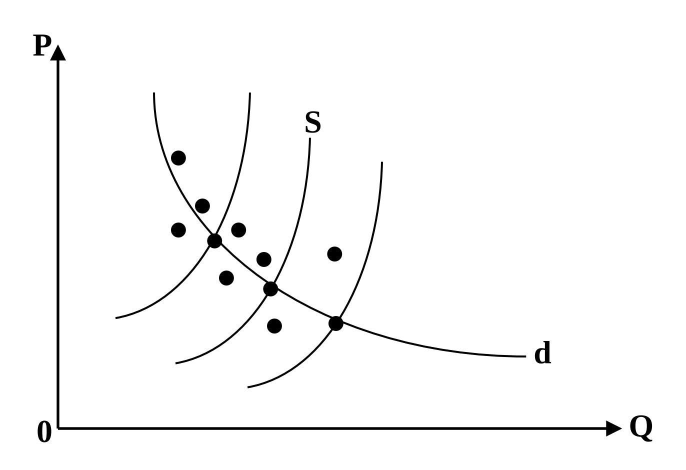
لأن منحنى الطلب ينتقل عبر الزمن فإن قيم التوازن لـ(P) و(Q) أيضاً تنتقل عبر الزمن . يجب أن يكون واضحاً من الشكل أن قيم التوازن تتبع مسار منحنى العرض ، وهكذا فإن منحنى العرض مشخصاً لأن معلمات العرض يمكن أن تستنتج من الشكل المختزل (الحركة في القيم التوازنية (P) و(Q) ، لاحظ ان الحركة التي يقوم بها (Y) عبر الزمن (أو عبر المشاهدات في تحليل المقطع العرضي) ضروري جداً لتشخيص معادلة العرض.

والان اذا أخذنا نموذج تكون فيه علاقة العرض مقررة بوساطة درجة الحرارة (T) في المنطقة ، و منحنى الطلب لا يتأثر بذلك ، عندئذ فإن معلومات مسبقة (a Priori information) حول متغير خارجي كان مستبعداً (درجة الحرارة) في معادلة العرض سوف يسمح لمنحنى الطلب أن يشخص ،

وطبعاً فإن منحنى الطلب ومنحنى العرض يمكن أن يشخصا ، والعلاقة الآتية للعرض والطلب تتضمن هذه الخاصية :

العرض : $Q_t = \alpha_1 + \alpha_2 P_t + \alpha_3 T + \varepsilon_t$

الطلب : $Q_t = \beta_1 + \beta_2 P_t + \beta_3 Y_t + u_t$



وكما ذكرنا سابقاً فإنه من المعادلات الهيكلية بإمكاننا أن نحل من أجل (M) من المتغيرات الداخلية ونشتق معادلات الشكل المختزل ومعلمات الشكل المختزل .

إن معادلة الشكل المختزل هي معادلة تعبر عن متغير داخلي على نحو وحيد بصيغ متغيرات مقررة مسبقاً ، وحد الإضطراب الإحتمالي (Stochastic Disturbance) ، ولتوضيح هذا لنأخذ النموذج الكينزي في تقرير الدخل :

دالة الإستهلاك : $C_t = \beta_0 + \beta_1 Y_t + U_t \qquad 0 \langle \beta_1 \langle 1$

متطابقة الدخل : $Y_t = C_t + I_t$

في هذا النموذج فإن (C) الإستهلاك و(Y) الدخل هما المتغيران الداخليان وإن (I) الإستثمار (الإنفاق الإستثماري) يعامل هنا بوصفه متغيراً خارجياً . إن كلا المعادلتين هيكليتين ، ومعادلة الدخل هي متطابقة ، وكما هو معتاد فإن الميل الحدي للإستهلاك (MPC) أو (β_1) إفترضت أنها تقع بين الصفر والواحد فإذا عوضنا دالة الإستهلاك في دالة الدخل فإننا نحصل على الشكل المختزل :

$$Y_t = \beta_0 + \beta_1 Y_t + U_t + I_t$$

$$Y_t - \beta_1 Y_t = \beta_0 + U_t + I_t$$

$$Y_t(1-\beta_1) = \beta_0 + U_t + I_t$$

$$Y_t = \frac{\beta_0}{1-\beta_1} + \frac{1}{1-\beta_1} I_t + \frac{U_t}{1-\beta_1}$$

$$Y_t = \underbrace{\frac{\beta_0}{1-\beta_1}}_{\pi_0} + \underbrace{\frac{1}{1-\beta_1}}_{\pi_1} I_t + W_t$$

$$Y_t = \pi_0 + \pi_1 I_t + W_t$$ معادلة الشكل المختزل (Y)

ومعادلة الشكل المختزل هذه ل(Y) تعرض المتغير الداخلي (Y) وحيداً بوصفه دالة في المتغير الخارجي (I) وحد الإضطراب الإحتمالي (U)و(π_1,π_0) هي معاملات الشكل المختزل المرافقة . يجب أن نلاحظ أن هذه المعاملات للشكل المختزل توليفات غير خطية للمعاملات أو المعلمات الهيكلية ، نعوض عن قيمة (Y) الموجودة في الشكل المختزل في معادلة (C) الأولى في النموذج :

$$C_t = \beta_0 + \beta_1(\pi_0 + \pi_1 I_t + W_t) + U_t$$

$$C_t = \beta_0 + \beta_1\left(\frac{\beta_0}{1-\beta_1} + \frac{1}{1-\beta_1} I_t + \frac{U_t}{1-\beta_1}\right) + U_t$$

$$C_t = \beta_0 + \frac{\beta_1\beta_0}{1-\beta_1} + \frac{\beta_1}{1-\beta_1} I_t + \frac{\beta_1 U_t}{1-\beta_1} + U_t$$

$$C_t = \frac{\beta_0(1-\beta_1) + \beta_1\beta_0 + \beta_1 I_t + \beta_1 U_t + U_t(1-\beta_1)}{1-\beta_1}$$

$$C_t = \frac{\beta_0 - \beta_1\beta_0 + \beta_1\beta_0 + \beta_1 I_t + \beta_1 U_t + U_t - \beta_1 U_t}{1-\beta_1}$$

$$Ct = \underbrace{\left(\frac{\beta_0}{1-\beta_1}\right)}_{\pi_2} + \underbrace{\left(\frac{\beta_1}{1-\beta_1}\right)}_{\pi_3} I_t + \underbrace{\left(\frac{U_t}{1-\beta_1}\right)}_{W_t}$$

$$C_t = \pi_2 + \pi_3 I_t + W_t$$

إن معاملات الشكل المختزل (π_1,π_2) تعرف أيضاً بوصفها أثر المضاعفات (Multipliers) ، وذلك لأنها تقيس التأثير على المتغير الداخلي المتأتي من تغيير بوحدة واحدة في المتغير الخارجي .

ففي النموذج الكنيزي السابق اذا ازداد الإنفاق الإستثماري (I) بمقدار دينار واحد ، وقد إفترضت (MPC) (الميل الحدي للإستهلاك) لتكون (0.8) عندئذ فإن $(\pi_1 = 5)$ ، وهذا يعني زيادة الإستثمار بمقدار دينار واحد سيقود في النهاية الى زيادة في الدخل (5) خمسة دنانير ، وهذه زيادة خمسة مرات ، وبالطريقة نفسها فإن $(\pi_1 = 4)$ وهذا يعني زيادة دينار واحد في الإستثمار سوف تقود الى (4) دينار في الإنفاق الإستهلاكي .

14-2 مشكلة التشخيص الإحصائي

يتصف النموذج الإقتصادي القياسي بثلاثة خصائص :

1- المحتوى الإقتصادي (Economic Content) للنموذج

2- الهيكل أو البناء الرياضي للنموذج (Mathematical Structure)

3- الصفات الإحصائية للنموذج (Statistical Properties)

إن الإتساق المنطقي (Logical Consistency) ، وكمال النموذج الإقتصادي القياسي تتقرر بالجانب الرياضي من النموذج ، أما الخصائص الإحصائية فهي تهتم بتقدير معلمات النموذج ، ومدى نجاح التقدير يعتمد على البيانات التجريبية وشكل النموذج ، فإذا لم يكن النموذج بالشكل الإحصائي الملائم ، فإن ذلك النموذج ينتج عنه معلمات ليس لها قيمة واحدة على الرغم من توفر البيانات الكافية ، وبلغة الإقتصاد القياسي فإن النموذج ربما لايكون مشخصاً .

أ- مشكلة التشخيص (The Problem of Identification)

لناخذ نموذج إقتصادي قياسي بسيط ساكن للسوق على سلعة واحدة (X) :

$$(1)..........X_d = a_0 + b_0P + u$$
$$(2)..........X_s = a_1 + b_1P + \upsilon$$
$$(3)..........X_d = X_s$$

وفي هذا النموذج فإن :

X_d = الكمية المطلوبة من السلعة X

X_s = الكمية المعروضة من السلعة X

P = سعر السلعة X

υ,u = متغيرات عشوائية مختلفة

b_1,b_0,a_1,a_0 = معلمات المجتمع الإحصائي وهي ثوابت .

فإذا كانت الثوابت (Constants) معروفة ، فإن المعـــادلات الهيكليـــة (Structural Equations) الثلاثة يمكن أن تحل للحصول الكميـــة المتبادلـــة وسعر التوازن إستناداً الى الأخطاء العشوائية (υ,u) ، إن السعر الفعلي والكمية الفعلية المتبادلة هي نتيجة لتفاعل دوال العرض والطلب .

بالنسبة الى قيم معطاة للمعلمات هناك زوج وحيد للمتغيرين (P) و(X) يتقرر ، وهذا يمكن أن يعبر عنه هندسياً بوصفه تقاطع منحنى الطلب مع منحنى العرض . إفترض إن بيانات عينة إحصائية تتألف من سلسلتين زمنيتين تبـــين سعر التوازن والكمية المتبادلة عبر كل فترة من الـــزمن والآن ينشـــأ الســـؤال التالي :

كيف نعرف أية معادلة يمكن أن تعبر عنها البيانات المشاهدة في السلسلة الزمنية ؟

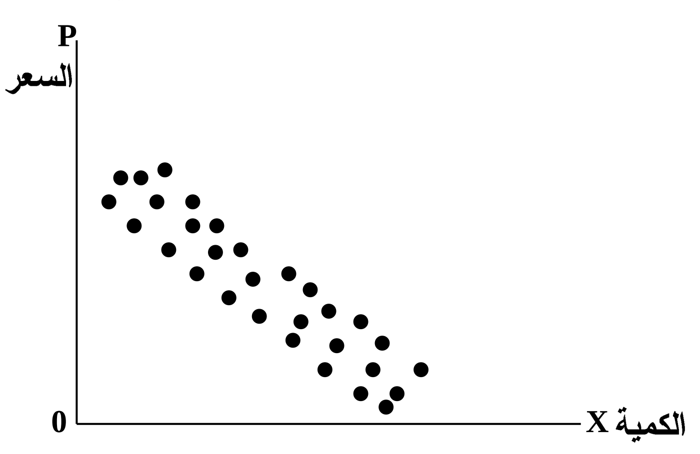
إن السلاسل الزمنية يمكن أن ترسم في شكل إنتشار ولكن كيف نستطيع أن نتأكد أن نقاط المشاهدة تقع على منحنى الطلب فقط أو على منحنى العرض فقط ؟ هذه الحالة هي مثال مشكلة التشخيص التي يجب أن نواجهها في نمـــاذج الإقتصاد القياسي ، في هذه الحالة يجب أن نسأل في مـــا اذا كانـــت البيانـــات تشخص منحنى الطلب أو منحنى العرض أو خليط منهما ؟

في بعض الأمثلة من السهولة نسبياً أن نقرر ماذا تبين البيانات ، في ظل الظروف الآتية نستطيع أن نصل الى إستنتاجات معينة :

1- من المعروف أن منحنى الطلب لم يتغير عبر الزمن ، وإن منحنى العرض ينتقل كثيراً ، فإذا كانت هذه هي الحالة فإن معلمات معادلة الطلب فـــي النموذج السابق هي في الحقيقة ثابتة ، ولكن معلمات معادلة العرض في الحقيقة قد تغيرت ، إن بيانات العينة الإحصائية سوف تبين نقاط منتشـــرة تقـــع فـــوق منحنى الطلب ، فإذا كان منحنى الطلب مستقراً نسبة الى منحنـــى العـــرض ،

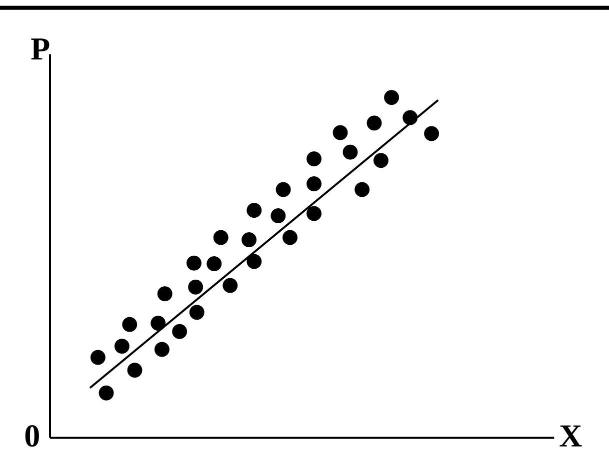
عندئذ فإن التغيرات في السعر والكمية المتاجر بها هي غالباً تعود الــى إنتقــال العرض .


إن شكل إنتشار العينة سوف يظهر مثل الشكل الآتي :



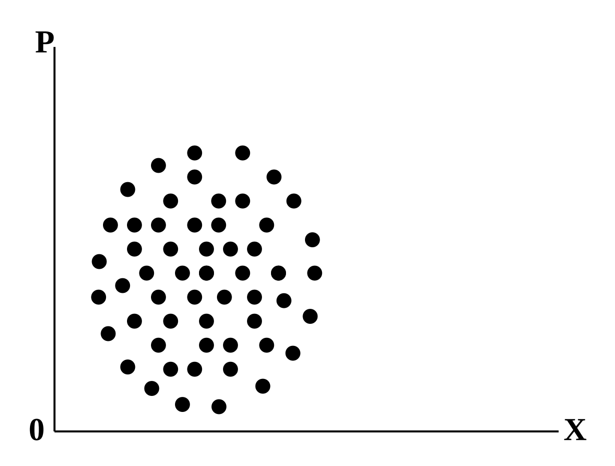
وبعد ذلك يمكن أن نقول أن منحنى الطلب شُخصْ مــن قبــل البيانات ، إن هذا الوضع تتصف به السوق للسلع الزراعية المتعددة عندما تكون تاثيرات سقوط الأمطار أو فترة الحصاد أو الأنماط الموسمية لإنتــاج الحليــب تسبب إنتقالات عنيفة في منحنى العرض ، وعندما يكون هناك دليلاً جيــداً أن معلمات منحنى الطلب (دخل المستهلك والأسعار الأخرى) قد تغيرت قليلاً جداً ، أو لم تتغير أبداً ، بينما معلمات منحنى الطلب قد تغيرت ، عندئذ فــإن معادلــة الطلب مثل المعادلة (1) السابقة يمكن أن تقدر بتحليل الإنحدار البسيط ، بمعادلة العرض مثل (2) لا يمكن أن تقدر .

2- من المعروف أن منحنى العرض لا يتغير بينما منحنى الطلب ينتقل على نحو كبير ، وهذه هي الحالة المعاكسة للحالة الأولى السابقة ، وهنا نقــول أنه من المعروف أن معلمات دالة العرض (اسعار عناصر الإنتــاج والحــالات التكنولوجية) تتغير قليلاً أو لا تتغير بتاتاً ، ولكن معلمات الطلب تتغير ، فــإن بيانات العينة الإحصائية سوف تبدو في الشكل الآتي :



والنقاط المشاهدة هنا تلحق (Trace Out) تقريباً منحنى العرض ، وهنا نقول أن منحنى العرض قد شخص أو أنه مشخصاً ، وفي هذا الموقف فإن الإنحدار البسيط يمكن أن يستخدم لتقدير معادلة العرض مثل المعادلة (2) السابقة ، ولكن دالة الطلب غير ممكن تقديرها .

3- من المعروف أن كلاً من منحنى الطلب ومنحنى العرض ينتقلان على نحو كبير ، وهذه الحالة هي المشتركة على الأغلب ، إن بيانات العينة الإحصائية تبين شكل الإنتشار الآتي :



عندئذ فإن الإنحدار البسيط للكمية المتاجر بها على سعر السوق لا تظهر معادلة الطلب أولاً أو معادلة العرض ، وإنما خليط مـــن كليهمـــا ، ونقـــول أن منحنى الطلب غير مشخص ومنحنى العرض أيضاً غير مشخص .

ب- معنى التشخيص الإحصائي

إن المناقشة السابقة يجب أن توفر لنا بعض التصورات في مـــا يتعلـــق بأهمية التشخيص ، ومع ذلك فإن التوضيح لايعد كاملاً بأية طريقة وذلك بسبب وجود ثلاثة إحتمالات مفتوحة :

الإحتمال الأول : هو ان النموذج ربما يكون تحت أو دون التشـــخيص (Underidentified) والذي فيه تكون مجموعة المعلمات الهيكلية غير ممكنـــة التقدير إحصائياً .

الإحتمال الممكن الثاني : إن النموذج ممكن أن يكون مشخصاً تمامـــاً ، وعندئذ فإنه من الممكن عموماً أن نحصل على قـــيم وحيـــدة لكـــل المعلمـــات للمعادلات الهيكلية (Structural Equations) .

الإحتمال الممكن الثالث : إن النموذج يمكن أن يكون فـــوق المشـــخص (Overidentified) ، وإن تقدير قيم وحيدة (Unique) للمعلمـــات الهيكليـــة عملية ممكنة فقط في ظل شروط مقيدة .

إن الإحتمالات الثلاثة يمكن أن توضع بالصيغة الآتية :

(1) دون التشخيص

(2) مشخصاً :

- مشخص تماماً
- فوق التشخيص

فإذا كانت معادلة واحدة من النموذج دون التشخيص أي غير مشخصـــة (Not Indentified) ، فإن النموذج يقال لـــه أنـــه دون التشـــخيص أو غيـــر مشخص ، وهكذا فإن النموذج المعروض في البحث السابق هو غير مشخص أو

دون التشخيص بغض النظر عن شكل الإنتشار الممكن . وبالطريقة نفسها فـإن النموذج يقال أنه فوق التشخيص (Overidentified) اذا كانت أية معادلة فيه فوق التشخيص ، وأن النموذج المشخص تماماً يتطلب أن كل معادلة تكون مشخصة .

إن صفة دون التشخيص أو صفة فوق التشخيص هي صفات نوعية مختلفة للنموذج ، فالأولى هي صفة غير إحتمالية ونقصد بذلك أن المعلمات الهيكلية يمكن تقديرها إحصائياً بغض النظر عن حجم ودقة البيانات في العينة الإحصائية ، ولذلك نركز هنا على الفرق بين نماذج دون التشخيص والنماذج المشخصة تماماً .

جـ قواعد التشخيص

لإغراض التشخيص يجب أن نستخدم نموذج السوق لسلعة واحدة ، ولكن يجب أن نتذكر أن المباديء التي يتم تطويرها هنا تطبق على نحو متساوي على بقية النماذج الإقتصادية القياسية :

بالنسبة للسلعة X لدينا :

(1) معادلة الطلب $X_d = a_0 + b_0P + u$

(2) معادلة العرض $X_s = a_1 + b_1P + \upsilon$

(3) معادلة التوازن $X_d = X_s$

وفي هذا النموذج فإن الرموز تعني المعاني نفسها في النموذج الأول الذي عرض في البداية ، يجب أن نسأل أنفسنا أولاً هل أن النموذج تام منطقياً ، وهذا يعني أو يماثل القول : اذا كانت قيم الثوابت (b_1, a_1, a_0, b_0) معروفة فهل يمكن للنموذج أن يحل للوصول الى القيم التوازنية للمتغيرات الداخلية (إعتماداً على (υ, u)) ؟ .

نلاحظ أن هناك ثلاثة متغيرات وثلاثة معادلات في هذا النموذج ، ولذلك فإن النموذج منطقياً تام .

فإذا أقنعنا أنفسنا بإن قيم المتغيرات يمكن أن تقـرر اذا كانـت قيمـة الثوابت معروفة ، فإننا بعد ذلك نسأل في ما اذا كان ممكناً تقـدير الثوابـت أو معلمات المجتمع الإحصائي إحصائياً ، وهذه هي مشكلة التشخيص ، قبل وضع الشروط التي يجب توافرها حتى يكـون النمـوذج مشخصـاً يجـب أن نعيـد التعريفات ، فالمعادلات الهيكلية هـي مـن نـوعين : المعـادلات السـلوكية (Behavioral Equations) وهي تحتوي على متغيرات نظامية ومتغيـرات عشوائية والتي تكـون فيهـا المتغيـرات النظاميـة أمـا متغيـرات داخليـة (Endogenous Variables) أو متغيـرات خارجيـة (Exogenous Variables) . وبقدر تعلق الأمر بالتشخيص ، فإننا يجـب أن نهتم بالمعادلات السلوكية التي تحتوي على متغيرات داخلية ، وفـي النمـوذج السابق فإن المعادلات (1) و (2) هي من هذا النوع ، وليس هنـاك متغيـرات خارجية في النموذج ، المعادلة (3) هي النوع الثاني وهي معادلـة تعريفيـة ، وبعبارة أكثر تحديداً هي معادلة تعريف التوازن .

ننتقل الآن الى قواعد التشخيص التي طورت من قبل أعضاء في لجنـة كاول (Cowles Comission at The University of Chicago) من أجل أن يكون النموذج مشخصاً من الضروري أن كل معادلة سلوكية تحتوي واحـداً أو أكثر من المتغيرات الداخلية مشخصة ، فإذا كانت أيـة واحـدة مـن هـذه المعادلات غير مشخصة فإننا نقول أن النموذج نفسه غير مشخص .

أولاً – شرط الدرجة للتشخيص (Order Condition For Identification)

إن هذا الشرط يعتمد على قاعدة حساب عـدد المتغيـرات الموجـودة والمتغيرات المستبعدة (Excluded) من معادلـة معينـة ، إن هـذا الشـرط ضروري ولكنه غير كاف لتشخيص معادلة معينة ، إن شرط الدرجة يمكـن أن ينص على ما يأتي :

من أجل أن تكون المعادلة مشخصة (أية معادلة) فإن العدد الكلي للمتغيرات (داخلية وخارجية) المستبعدة من المعادلة يجب أن يساوي أو أكبر من عدد المتغيرات الداخلية في النموذج ناقصاً واحد ، ومعلوم انه في النموذج التام فإن عدد المتغيرات الداخلية يساوي عدد المعادلات في النموذج ، بعبارة أخرى فإن شرط الدرجة في بعض الأحيان يكتب بصيغة مساوية أخرى :

من أجل أن تكون أية معادلة مشخصة فإن العدد الكلي للمتغيرات المستبعدة (Excluded) منها والمتضمنة أو الموجودة في بقية المعادلات يجب على الأقل أن يساوي عدد المعادلات في نظام المعادلات ناقصاً واحد .

لنجعل :

G = العدد الكلي للمعادلات = العدد الكلي للمتغيرات الداخلي

K = العدد الكلي للمتغيرات في النموذج (الداخلية والمقررة مسبقاً)

M = عدد المتغيرات (الداخلية والخارجية) الموجودة في معادلة معينة

عندئذ فإن شرط الدرجة للتشخيص يمكن أن يكتب بصيغة الرموز :

$$(K - M) \geq (G - 1)$$

المتغيرات المستبعدة ≤ **العدد الكلي للمعادلات ناقصاً واحد**

مثال : اذا كان هناك نظام يحتوي على (10) معادلات مع متغيرات عددها (15) ، عشرة منها متغيرات داخلية وخمسة متغيرات خارجية ، فإن معادلة تحتوي على (11) متغيراً تكون غير مشخصة ، بينما معادلة أخرى تحتوي (5) متغيرات تكون مشخصة .

أ- بالنسبة للمعادلة الأولى لدينا :

$K=15 \quad G=10 \quad M=11$

فإن شرط الدرجة :

$$(K-m) \geq (G-1)$$
$$(15-11)<(10-1)$$

وهذا يعني أن شرط الدرجة غير متحقق ، وأن المعادلة دون التشخيص .

ب- بالنسبة للمعادلة الثانية لدينا :

$K=15$ 	$G=10$ 	$M=5$

شرط الدرجة :

$$(K-m) \geq (G-1)$$
$$(15-5) <(10-1)$$

وهذا يعني أن شرط الدرجة متحقق

ثانياً – شرط الرتبة (The Rank condition For Indentification)

ينص شرط الرتبة على أنه :

في أي نظام متكون من (G) من المعادلات فإن أية معادلة معينة تكون مشخصة اذا كان من الممكن بناء على الأقل محددة واحدة (One Determinant) لاتساوي صفر ومن رتبة (G-1) من معاملات المتغيرات المستبعدة من معادلة معينة ولكنها موجودة أو محتواة في بقية معادلات النموذج .

إن الخطوات المحددة لمتابعة تشخيص معادلة من معادلات نموذج هيكلي (Structural Model) يمكن أن ندرجها كما يأتي :

أولاً : نكتب معلمات (Parameters) كل معادلات النموذج في جدول منفصل ، آخذين بنظر الإعتبار أن معلمة المتغير المستبعد من معادلة معينة مساوية للصفر .

مثلاً : لنجعل النموذج الهيكلي هو :

$$y_1 = 3y_2 - 2X_1 + X_2 + U_1$$
$$y_2 = y_3 - X_3 + U_2$$
$$y_3 = y_1 - y_2 + 2X_3 + U_3$$

عندما يكون (y_1,y_2,y_3) هي المتغيرات الداخليـــة ، وإن (X_1,X_2,X_3) هي المتغيرات المقررة مسبقاً ، إن هذا النموذج يمكن إعادة كتابتـــه بالصـــيغة الآتية :

$$-y_1 + 3y_2 + Oy_3 - 2X_1 + X_2 + OX_3 + U1 = 0$$
$$Oy_1 - y_2 + y_3 + OX_1 + OX_2 + X_3 + U2 = 0$$
$$y_1 - y_2 - y_3 + OX_1 + OX_2 - 2X_3 + U3 = 0$$

فإذا تجاهلنا المتغيرات العشـــوائية (u_3,u_2,u_1) فـــإن جـــدول معلمـــات النموذج هو كما يأتي :

المعادلات Equations	المتغيرات (Variables)					
	Y_1	Y_2	Y_3	X_1	X_2	X_3
المعادلة الأولى	−1	3	0	−2	1	0
المعادلة الثانية	0	−1	1	0	0	1
المعادلة الثالثة	1	−1	−1	0	0	−2

ثانياً : نشطب الصف من المعلمات للمعادلة التي يراد تشخيصها ، مثلاً اذا اردنا أن نختبر إمكانية تشخيص المعادلة الثانية من النموذج فإننـــا نشـــطب الصف الثاني من جدول المعلمات .

ثالثاً : نشطب الإعمدة التي فيها قيم معلمات لا تســـاوي صـــفراً فـــي المعادلة التي نختبرها لأغراض التشخيص .

إن حذف الصف الملائم والأعمدة الملائمة يجعلنا نبقى مـــع معلمـــات لمتغيرات غير موجودة أو متضمنة في المعادلة المعينة (المعادلة قيد التشخيص) ولكنها موجودة في بقية معادلات النموذج .

مثلاً اذا كنا نختبر تشخيص المعادلة الثانية من النظام ، فإننـــا ســوف نشطب العمود الثاني والعمود الثالث والعمود السادس من الجدول أعلاه وهكــذا نحصل على الجدولين الآتيين :

	Y1	X1	X2
المعادلة الأولى	**−1**	**−2**	**1**
المعادلة الثانية	**1**	**0**	**0**
المعادلة الثالثة			

	Y_1	Y_2	Y_3	X_1	X_2	X_3
المعادلة الأولى	**−1**	**3**	**0**	**−2**	**1**	**0**
المعادلة الثانية	**0**	**−1**	**1**	**0**	**0**	**1**
المعادلة الثالثة	**1**	**−1**	**−1**	**0**	**0**	**2−**

رابعاً : من المحددات من رتبة (G-1) وتفحص قيمهــا : فــإذا كانــت واحدة على الأقل من المحددات لا تساوي صفراً فإن المعادلة تكون مشخصة ، واذا كانت كل المحددات من رتبة (G-1) تساوي صفراً فإن المعادلة تحــت أو دون التشخيص .

في المثال السابق لإستكشاف تشخيص المعادلة الثانية مــن المعــادلات الهيكلية فإننا نحصل على ثلاثة محددات من رتبة (G-1) = 3−1 = 2 وهي :

$$\Delta1=\begin{vmatrix} -1 & -2 \\ 1 & 0 \end{vmatrix} \neq 0 \qquad \Delta2=\begin{vmatrix} -2 & 1 \\ 0 & 0 \end{vmatrix} =0 \qquad \Delta3=\begin{vmatrix} -1 & 1 \\ 1 & 0 \end{vmatrix} \neq 0$$

إن الرمز (Δ) يمثل المحددة (Determinant) نلاحظ أننا نســتطيع أن نشكل محددتين من درجة $\begin{bmatrix} (G-1) \\ 3-1=2 \end{bmatrix}$ لا تساوي صفراً ، وهكذا فالمعادلة الثانية من النموذج مشخصة .

خامساً : ومن أجل أن نرى في ما اذا كانت المعادلة مشخصـــة تمامـــاً (Exactly Identified) أو أنها فوق التشـــخيص (Overidentified) فإننـــا نستخدم شرط الدرجة (K-m) ≥ (G-1) مع المعيار (Criteria) الآتي :

اذا تحققت إشارة المساواة أي اذا كان :

$$(K-m) \geq (G-1)$$

فإن المعادلة مشخصة تماماً .

واذا تحققت إشارة عدم المساواة (Inequality) أي اذا كان :

$$(K-m) > (G-1)$$

فإن المعادلة فوق التشخيص (Overidentified) .

في حالة المعادلة الثانية السابقة فإننا نجد أن :

$$K=6 \qquad G=3 \qquad M=3$$

$$(K-m) \geq (G-1)$$

$$(6-3)>(3-1)$$

ولذلك فإن المعادلة الثانية من النمـــوذج هـــي فـــوق التشـــخيص ، إن تشخيص دالة معينة يتم أو ينجز من خلال إفتراض أن بعض المتغيـــرات فـــي النموذج لها معلمة أو معامل يساوي صفر في هذه المعادلة ، بمعنى أننا نفترض أن بعض المتغيرات لا تؤثر مباشرة على المتغير المعتمد في هـــذه المعادلـــة .

وهذا على كل حال إفتراض يمكن أن يختبر مع بيانات عينة إحصائية .

مثال : نموذج يصف سوق احد المنتوجات الزراعية :

$$D = a_0 + a_1P_1 + a_2P_2 + a_3Y + a_4t + u$$
$$S = \beta_0 + \beta_1P_1 + \beta_2P_2 + \beta_3C + \beta_4t + W$$
$$D = S$$

عندما يكون :

D = الكمية المطلوبة

S = الكمية المعروضة

P1 = سعر سلعة معطى

P2 = أسعار بقية السلع

Y = الدخل

C = الكلفة (الرقم القياسي لأسعار عناصر الإنتاج)

t = الإتجاه الزمني

في دالة الطلب يمثل الأذواق في دالة العرض يمثل التكنولوجيا .

شرط الدرجة (Order Condition)

$$\mathbf{(K-m) \geq (G-1)}$$

وفي مثالنا نجد أن :

K=7 G=3 M=5

ولذلك فإن :

$$(K-m) = (G-1)$$
$$(7-5)=(3-1)$$
$$2=2$$

وهكذا فإن المعادلة الثانية تحقق الشرط الأول الضروري للتشخيص .

شرط الرتبة (Rank Condition)

إن جدول المعلمات الهيكلية للنموذج هو كما يأتي :

المعادلات Equations	المتغيرات (Variables)						
	D	P_1	P_2	Y	t	S	C
المعادلة الأولى	-1	a_1	a_2	a_3	a_4	0	0
المعادلة الثانية	0	β_1	β_2	0	β_4	-1	β_3
المعادلة الثالثة	-1	0	0	0	0	1	0

1−	a3
−1	0

	−1	d3	
Δ =			= (0)(-1) - (-1)(a3) = a3
	−1	0	

إن قيمة المحددة لا تساوي صفر بمعنى أن ($a_3 \neq 0$) وهكـــذا نـــرى أن شرط الدرجة وشرط الرتبة قد تحققا ولذلك فإن المعادلة الثانية مـــن النمـــوذج مشخصة ، وإلا بعد من هذا فإننا نرى أنه في شرط الدرجة فإن إشارة المساواة تحققت :

$$7 - 5 = 3 - 1 = 2$$

وبالنتيجة فـــإن المعادلـــة الهيكليـــة الثانيـــة هـــي مشخصـــة تمامـــاً (Exactly Identified) .

الفصل الخامس عشر

تطبيقات وتمارين محلولة

Applications and Solved Exercises

تطبيقات وتمارين محلولة

15-1 بعض التمارين المحلولة

1- إفترض أنك أعطيت نموذج الإنحدار الآتي :

$$\hat{Y}_t = \beta_0 + \beta X_t + U_t$$

عندما تكون المشاهدات حول (Y,X) هي

4	7	5	6	8	$=Y_t$
2	1	3	2	4	$=X_t$

المطلوب : قدر المعلمات (β_0 و β_1) وقدر الإنحراف المعياري للمتغير العشوائي .

الحل : المعادلات الطبيعية لهذا النموذج هي :

$$\Sigma Y_t = n\hat{\beta}_0 + \hat{\beta}_1 \Sigma X_t$$

$$\Sigma X_t Y_t = \hat{\beta}_0 \Sigma X_t + \hat{\beta}_1 \Sigma X_t^2$$

الحسابات التي يمكن أن نجريها هي :

$$\sum_{t=1}^{5} X_2^2 = 34 \qquad \sum_{t=1}^{5} X_t = 12 \qquad \sum_{t=1}^{5} Y_t = 30$$

$$\sum_{t=1}^{5} X_t Y_t = 74 \qquad N = 5$$

نطبق النتائج على المعادلتين الطبيعيتين :

$$30 = 5\hat{\beta}_0 + 12\hat{\beta}_1$$

$$74 = 12\hat{\beta}_0 + 34\hat{\beta}_1$$

وبحل هاتين المعادلتين نحصل على :

$$\hat{\beta}_0 = 5.075$$

$$\hat{\beta}_1 = 0.385$$

أما الإنحراف المعياري للمتغير العشوائي :

$$\hat{\delta}_u^2 = \frac{\sum\left(Y_t - \overline{Y}\right)^2}{n-2} = 3.077$$

2- يوضح أحد محللي الإستهلاك إن دالة الإستهلاك :

$$C_t = \beta_0 + \beta_1 Y_t$$

غير مفيدة لأن النقاط (C_t, Y_t) في شكل الإنتشار لا تقـع علـى خـط مستقيم ، ويلاحظ ذلك المحلل أيضاً أنه في بعض الأحيان يرتفـع (Y_t) ولكـن (C_t) ينخفض ، ولذلك فهو يستنتج أن (C_t) هي ليست دالة في الـدخل (Y_t) ، هل أن تحليله صحيحاً أم لا ؟ ولماذا ؟

الحل : إن تحليله غير صحيح لأنه أهمل صيغة الخطأ العشـوائي (U_t) والتي تأخذ قيماً سالبة وقيماً موجبة ، ولكن قيمته المتوقعة $E(U_t)$ هي صفراً .

إن العلاقة $(C_t = \beta_0 + \beta_1 Y_t)$ هي ليست علاقة تامة ولكنها علاقة على مستوى الوسط الحسابي أو هي علاقة إحتمالية .

3- اذا أعطيت المتغيرات (X_3, X_2, X_1) وكان تباين هذه التغيرات كما يأتي :

$$\delta_3^2 = 5.0 \qquad \delta_2^2 = 3.0 \qquad \delta_1^2 = 1$$

فإذا علمت أن هذه المتغيرات تفسـيرية أو توضـيحية وإن (Y) هـو المتغير المعتمد بحيث قدرت الدالة الآتية :

$$\hat{Y} = 13 - 2X_1 - 10X_3$$

المطلوب : أحسب تباين المتغير المعتمد (Y)

الحل : نستخدم علاقة أساسية للتباين لمجموع المتغيرات الإحتمالية وهذه العلاقة تنص على ما يأتي :

اذا كانت : $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_n X_n$

والمتغيرات $(X_n.......X_1)$ هي متغيرات تفسيرية أو توضيحية عندئذ فإن :

$$Y \text{ تباين } = \beta_1^2\delta_{x1}^2 + \beta_2^2\delta_{x2}^2 + \dots\dots\dots + \beta_n^2\delta_{xn}^2$$

فإذا طبقنا هذه الصيغة على السؤال الحالي نحصل على :

$$Var(Y) = 4+27+500$$
$$= 531$$

4- يقال أن العوائل ذات الدخل المتوسط والعوائل ذات الدخل العالي في الدول المتقدمة تغادر مراكز المدن لإن الضرائب في المدن أعلى من الضرائب في مناطق الضواحي التابعة للمدن . إفترض أن لدينا بيانات تتعلق بعدد من المدن في نقطة محددة من الزمن . ضع فرضية لهذه المشكلة أو الظاهرة بصيغة نموذج إنحدار .

أ- يمكن وضع نموذج إنحدار لتلك الفرضية كما يأتي :

$$Y_i = \beta_0 + \beta_1(T_{ci} - T_{si}) + U_i$$

عندما يكون :

معدل دخل العائلة في المدن = Y_i

معدل الضريبة في المدن = T_{ci}

معدل الضريبة في ضواحي المدن = T_{si}

حد الإضطراب (المتغير العشوائي) = U_i

ب- يمكن وضع نموذج إنحدار آخر كما يأتي :

$$Y_i = \beta_0 + \beta_1 \frac{T_{ci}}{T_{si}} + U_i$$

وفي كلا الوضعين فإن قيمة (β_1) تكون أقل من الصفر (سالبة) فإذا كان (T_c) أعلى نسبياً من (T_s) سوف نتوقع أن العوائل ذات الدخل المتوسط والدخل العالي تعيش في الضواحي .

5- لنأخذ النموذج الآتي :

$$Y_t = \beta_0 + \beta_1 X_t + U_t \dots\dots\dots\dots(1)$$

وفي هذه المعادلة (1) نجد أن حد الإضطراب أو المتغيــر العشــوائي (U_t) يعتمد على المتغير المستقل (X) بالنمط الآتي :

$$U_t = \beta_2 + \beta_3 X_t + \varepsilon_t \text{...........(2)}$$

عندما يكون (ε_t) هو المتغير العشوائي أو حد الإضطراب غير مستقل عن (X_t) ، والذي تنطبق عليه كل الإفتراضات الموضوعة لتحليل الإنحدار .

إفترض أن (β_3) أكبر من الصفر ($\beta_3>0$) ، بين أو وضـــح إن (β_1) في المعادلة (1) تقلل من تأثير (X_t) على (Y_t) .

الحل : نعوض المعادلة (2) في المعادلة (1) فنحصل على :

$$Y_t = (\beta_0 + \beta_2) + (\beta_1 + \beta_3)X_t + \varepsilon_t$$

6- لماذا يكون الوسط الحسابي المقدر من عينة إحصائية متكونة مـــن (30) مشاهدة أفضل من الوسط الحسابي المقدر من عينة إحصائية متكونة مـــن (20) مشاهدة ، وهل ان كلا الوسطين المقدرين غير متحيزين ؟

الحل : ليكون الوسط الحسابي لعينة الـــ 20 مشاهدة = ($\overline{X}_{20}$)

ليكون الوسط الحسابي لعينة الـــ 30 مشاهدة = ($\overline{X}_{30}$)

إعتماداً على حجم العينة الإحصائية ، وعندئذ اذا كانت كلا العينتين قـــد سحبت من المجتمع الإحصائي نفسه فإن :

$$E(\overline{X}_{20}) = E(\overline{X}_{30}) = M$$

عندما يكون (M) هو الوسط الحسابي للمجتمع الإحصائي .

وهكذا فإن كلا الوسطين المقدرين غير متحيزين ولكن ($\overline{X}_{30}$) ســـوف يكون مفضلاً على الآخر لأن تباين ($\overline{X}_{30}$) سيكون أصغر من تبـــاين ($\overline{X}_{20}$) ، ولذلك فإن إستخدام ($\overline{X}_{30}$) (العينة ذات 30 مشاهدة) سيقود الى حدود ثقة أضيق وإختبارات للفرضيات يمكن الإعتماد عليها .

7- لنأخذ وضعاً يكون فيه عدد من الشركات موجوداً في مدينة معينـــة يعتمد على معدل الضريبة النسبي ، إفترض أيضاً أنه على الرغم من أن فوائـــد

ضريبية ربما تحصل للمقيمين في المدينة كلما كان عدد المقيمين هناك أكثـــر ، ولكن معدل التلوث يكون أو سيكون أكبر . عن موقع الشركة وعلاقات التلوث بصيغ نماذج الإنحدار .

الحل : لنجعل (N_t) = عدد الشركات المتواجدة في تلك المدينة

عندئذ فإن نموذج موقع الشركة ربما يعرض كما يأتي :

$$N_t = \beta_0 + \beta_1 (\frac{T_{1t}}{T_{2t}}) + U_{1t}$$

عندما يكون :

(T_{1t}) = الضريبة في تلك المدينة

(T_{2t}) = متوسط معدل الضريبة في المدن المجاورة

نحن نتوقع أن (β_1) أصغر من الصفر ، وبالطريقة نفسها لنجعـــل (P_t) مقياساً للتلوث ، وعندئذ علاقة التلوث ربما تعرض كما يأتي :

$$P_t = \beta_2 + \beta_3 N_t + U_{2t}$$

وهنا نتوقع أن (β_3) أكبر من الصفر .

8- اذا أعطيت دالة الإنتاج الآتية :

$$Q_t = (\frac{1}{A}) L_t^{\beta 0} K_t^{\beta 1} \ell^{ut}$$

عندما يكون :

(Q) = الإنتاج

(t) = الزمن

(L_t) = العمل

(U_t) = حد الإضطراب أو المتغير العشوائي

(K_t) = رأس المال

إفترض أن القيمـــة المتوقعـــة للمتغيـــر العشـــوائي $(E(U_t)=0)$ وأن $(E(U_t^2)=\delta^2)$ وأن (U_t) مستقل عن (K_t) (L_t) .

المطلوب : إقترح طريقة لتقدير المعلمات$(A،\beta_0،\beta_1)$.

الحل :

بإستخدام التحويل اللوغاريتمي نحصل على ما يأتي :

$$Q_t^1 = \beta + \beta_0 L_t^1 + \beta_1 K_t^1 + U_t$$

عندما يكون :

$$Q_t^1 = Log_e Q_t$$

$$B = \log_e(\frac{1}{A})$$

$$L_t^1 = Log_e L_t$$

$$K_t^1 = Log_e K_t$$

ثم نقدر $(B،\beta_0،\beta_1)$ ونأخذ $(\hat{A} = \ell^{-\hat{\beta}})$

9- إفترض إن المصروفات الإستثمارية لشركة معينة تعتمد على معدل سعر الفائدة ومعدل الأرباح والتغير في المبيعات بوصفها مؤشراً للتوقعات .

أ- أكتب نموذج لهذه العلاقة (نموذج إنحدار)

ب- إفترض إن الأرباح خلال فترة أو مدة العينة الإحصـــائية بالنســـبة للشركة كانت في المدة 15% . ناقش أية مشكلات تنشـــأ فـــي عمليـــة تقـــدير النموذج .

الحل :

أ- يكون نموذج الإنحدار كما يأتي :

$$I_t = \beta_0 + \beta_1 r_t + \beta_2 \pi_t + \beta_3 \Delta S_t + U_t$$

عندما يكون :

(I_t) = المصروفات الإستثمارية

(π_t) = معدل الربح

(ΔS_t) = التغيير في المبيعات

(r_t) = معدل سعر الفائدة

(U_t) = حد الإضطراب أو المتغير العشوائي

ب- إن المشكلة في تقدير هذا النموذج سببها التعـدد الخطـي التـام ، وبصيغة أكثر تحديداً نقول اذا كان معدل الربح 15% لكل فترة زمنيـة فإننـا لايمكن أن نقدر المعلمة $(\hat{\beta}_0)$ والمعلمة $(\hat{\beta}_1)$.

10- اذا أعطيت النموذج الآتي :

$$C_t = \beta_0 + \beta_1 C_t + \beta_2 Y_t + \varepsilon_{1t} \ldots\ldots\ldots\ldots\ldots(1)$$

$$Y_t = I_t + C_t \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(2)$$

$$I_t = a_0 + a_1 Y_t + a_2 Y_{t-1} + a_3 r_t + \varepsilon_{2t} \ldots\ldots(3)$$

عنما تكون :

(C) = المصروفات الإستهلاكية

(I) = الإستثمار

(Y) = الدخل

(r) = معدل سعر الفائدة

إفترض أن (ε_1) (ε_2) غير مرتبطان ذاتياً ومستقلان عن (r_t) .

أ- حدد المتغيرات الداخلية والمتغيرات الخارجية في النموذج

ب- كيف ستقدر المعادلة (1) ؟

جـ كيف ستقدر المعادلة (2) ؟

الحل :

أ- المتغيرات الداخلية هي : (I_t) (Y_t) (C_t)

المتغيرات الخاجية هي : (r_t) (Y_{t-1}) (C_{t-1})

ب- إن المشكلة في المعادلة (1) هي أن (Y_t) مرتبطة بـ(ε_{1t}) ولذلك فإن الإجراء الصحيح هو أن تحل $(\hat{Y}_t)$ محل (Y_t) ، نحصل على $(\hat{Y}_t)$ من خلال إنحدار (Y_t) على كل المتغيرات الخارجية التي هي (C_{t-1}) (Y_{t-1}) (r_t) وهكذا نجد أن $(\hat{Y}_t)$ ستكون :

$$\hat{Y}_t = \hat{\gamma}_0 + \hat{\gamma}_1 C_{t-1} + \hat{\gamma}_2 Y_{t-1} + \hat{\gamma}_3 r_t + U_t$$

تقدر المعادلة (1) في النموذج بالطريقة الإعتيادية بعد إحلال $(\hat{Y}_t)$ محل (Y_t) أي أن المعادلات الطبيعية يمكن الحصول عليها من خلال المساواة للصفر القيم :

$$\sum \varepsilon_1^* = 0 \qquad \sum(\varepsilon_1^*, C_{t-1}) = 0 \qquad \sum(\hat{\varepsilon}_1^*, \hat{Y}_t) = 0$$

15-2 بعض التطبيقات الإقتصادية القياسية

الحالة الأولى :

دراسة العلاقة بين الدخل الإجمالي والضرائب من خلال البيانات حول (51) ولاية أمريكية ، إن كلاً من الدخل والضرائب تقاس ببلايين الدولارات ، وتم تقدير النموذج وإعطاء النتائج مع قيم إختبار (t) وضعت بين الأقواس كما يأتي :

$$T\hat{a}x = -0.305 + 0.14492$$
$$(-3.04) \quad (121.7)$$

$$R^2 = 0.997 \qquad d.f. = 49 \qquad F = 14.820 \qquad \hat{\delta} = 0.549$$

إن هذا النموذج يعطي توفيقاً جيداً جداً وذلك لأن 99.7 بالمائة من الإنحرافات في الضرائب موضح بوساطة الدخل . كما أن قيمة إختبار (F)

عالية جداً والتي تشير الى علاقة قريبة أو قوية ، إن قيمة إختبار (t) للمعلمتين ($\hat{\alpha}$ و $\hat{\beta}$) ذات قيمة معنوية عند مستوى معنوية 5% .

إن التأثير الحدي (Marginal Effect) للدخل علـــى الضـــرائب هـــو (0.14492) ، وهذا يعني أن زيادة (100) دولار في الدخل من المتوقع أن تزيد الضرائب على مستوى المعدل بمقدار (14.49) دولار .

الحالة الثانية :

في هذه الحالة نعرض نموذج إنحدار خطي بسيط قد أخذ من البحـــوث المنشورة . ففي دراسة لكل من (Bell) ومورفي (Murphy) في (1969) ، تم تقدير هيكل الكلفة لشركات الصيرفة التجارية بإســـتعمال بيانـــات (SMSAs) (Standard Metropolition Statistical Areas) في شمال غرب الولايات المتحدة الأمريكية ، وهنا نكتفي بجزء من تلك الدراسة .

إن البيانات هي من نوع بيانات المقطع العرضـــي لــــ(41 SMSAs) لسنة (1965) ، في المرحلة الأولى من التحليل الذي عرضته هذه الدراسة تـــم تقدير العلاقة بين تكاليف العمليات المباشرة للمصارف وعدد مـــن المتغيـــرات التوضيحية :

عدد الودائع تحت الطلب (demand deposit) ، سعر الفائـــدة علـــى رأس المال ، الأجور ، أسعار المواد المستعملة في حســـابات ودائـــع الطلـــب وغيرها من المتغيرات التوضيحية ، إن النموذج المستعمل هنا هو كما يأتي :

$$\mathbf{AC_t = AN_t^{II}U_t}$$

عندما يكون :

AC_t = معدل الكلفة المقدر لــ(t^{th}) مصرف

N_t = عدد حسابات ودائع الطلب

U_t = حد الإضطراب

(II)(A) هي معلمات (Parameters) مجهولة مطلوب تقديرها ، وبأخذ لوغاريتم جانبي المعادلة السابقة فنحصل على :

$$\mathbf{lnAC_t = lnA + IIlnN_t + lnU_t}$$
$$= \alpha + \beta\, \mathbf{lnN_t + U_t}$$

وهذه سوف يوصفها نموذج لوغاريتمي مزدوج .

وفي الجدول الآتي نضع بعض النتائج التي حصلت من هذا النموذج ولمدن مختارة في الولايات المتحدة الأمريكية :

قيمة t لمعلمة II					
المدينة أو المنطقة	Log A	II	t- Value for II	R^2	درجات الحرية
Rochester	3.3614	0.0044	0.6783	0.0280	16
Philadelphia	3.2861	0.0077	2.6291	0.0834	76
Syracuse	3.3177	0.0037	0.9602	0.1033	8
New Britain	3.7693	−0.0146	−4.3014	0.8605	3
New York	3.2795	0.0130	4.5137	0.2114	76
Buffalo	3.1543	0.0237	3.5389	0.6415	7

نلاحظ من الجدول أن عدداً قليلاً من (R^2) ذات قيمة عالية ، إن نتائج الدراسة تشير الى أن وفورات الحجم (Economies of Scale) التي يمكن أن تتحقق من قبل المصارف الكبيرة يقابلها أو يعوض عنها (offset by) بإيجاد الفروع (branching) . وهكذا فإن المصارف الصغيرة يمكنها أن تعمل إقتصادياً (Economically) كما تعمل المصارف الكبيرة ذات الفروع الصغيرة العديدة أو الكثيرة .

الحالة الثالثة :

مثال حول مشكلة الإرتباط الخطي المتعدد (Multicollinearity) ، في هذا المثال نربط بين الإنفاق المتراكم على إدامة السيارة وعمر السيارة وعدد الأميال المقطوعة ، لنجعل (E_t) الإنفاق المتراكم على إدامة السيارة في الزمن

(t) (بإستثناء البنزين) لسيارة معينة (X_1) عدد الأميال المتراكمة بآلاف الأميال و(X_2) العمر (عمر السيارة) بالأسابيع منذ الشراء الأصلي ، وهنا يمكن ان يكون لدينا النماذج البديلة الآتية :

النموذج الأول : $Et = \alpha_0 + \alpha_1 X_2 + U_1 t$

النموذج الثاني : $Et = \beta_0 + \beta_1 X_1 + U_2 t$

النموذج الثالث : $Et = \nu_0 + \nu_1 X_2 + \nu_2 X_1 + U_3 t$

فالسيارة التي يتم سياقتها كثيراً سيكون لها مصاريف إدامه أكبر ، وبالطريقة نفسها فإن السيارة نفسها كلما كانت قديمة فإن تكاليف إدامتها تكون أكبر ، وأيضاً فإن السيارة التي لها عدد أميال مقطوعة أكبر من المحتمل جداً أن يكون لها إنفاق على الإدامة أكبر ، ولذلك تتوقع ان المعلمات (α_1 و β_1 و γ_1 و γ_2) أن تكون موجبة الإشارة وفي ما يأتي نتائج تقدير المعلمات وقيم إختبار (t) المرافقة لتلك النماذج الثلاثة :

النموذج الأول

$$\hat{E}_t = -626.24 + 7.35 X_2$$
$$(-5.98) \quad (22.16)$$

$\overline{R}^2 = 0.89$
درجات الحرية = 55

النموذج الثاني

$$\hat{E}_t = -796.07 + 53.45 X_1$$
$$(-5.91) \quad (18.27)$$

$\overline{R}^2 = 0.85$
درجات الحرية = 55

النموذج الثالث

$$\hat{E}_t = 7.29 - 151.5 X_1 + 27.58 X_2$$
$$(0.06) \quad (-7.06) \quad (9.58)$$

$\overline{R}^2 = 0.94$
درجات الحرية = 54

ثم تقدير هذه النماذج بإستعمال بيانات فعلية تخص سيارة تويوتا (1971) صالون وهي معروضة في الجدول الاتي :

جدول يبين بيانات عن كلفة إدامة سيارة تويوتا 1971 وعدد الأميال المقطوعة (X_1) وعمر تلك السيارة بالأسابيع (X_2)

X_2	X_1	الكلفة	X_2	X_1	الكلفة
5	0.8	11	279	43.7	1182
12	3.0	16	281	44.3	1231
30	4.9	55	313	47.6	1244
40	7.1	66	326	48.9	1257
42	7.6	76	328	49.1	1260
53	10.1	83	329	49.2	1342
66	12.0	135	336.5	50.0	1356
73	12.8	160	338	50.1	1467
79	13.9	163	342.5	50.6	1518
101	18.6	211	344.5	50.8	1557
114	21.1	258	351	51.6	1565
129	23.2	322	366	53.2	1583
150	25.3	374	384	55.7	1609
180	28.7	408	388	56.0	2825
195	30.5	478	402	57.3	2893
196	30.6	489	432	60.2	2918
204	31.4	536	433	60.3	3011
212	32.9	590	436	60.6	3077
224	35.3	604	436	63.0	3095
227	35.3	704	456	63.7	3154
232	36.6	985	463.5	63.9	3162
235	37.0	1021	465	65.1	3217
239	38.1	1030	478	65.8	3274
249	39.5	1096	485	67.7	3320
260	40.7	1114	498.5	72.1	3329
271	43.0	1134	526	72.1	3401
272	43.1	1157	527	73.6	3412
275.5	43.2	1176	538	74.4	3425
276	43.4	1182			

من الملاحظ ان معلمة المتغير (X_1) موجبة الإشارة في النموذج الثاني فإنها سالبة الإشارة على نحو مهم إحصائياً في النموذج الثالث ، وهكذا فإن هناك تعاكس او تضارب في الإشارة ، أما معلمة المتغير (X_2) قد تغيرت على نحـــو كبير ومهم ، كما أن قيمة إحصاءة (t) للمتغير (X_2) والمتغير (X_1) هي أقـــل كثيراً في النموذج الثالث ، وهنا فإن السبب لوجود هذا التغيير المعنوي إحصائياً في النتائج هـــو الإرتبـــاط العـــالي (High Correlation) بـــين المتغيـــرين التوضيحيين (X_2،X_1)وفي هذه الحالة فإن الإرتباط بينهما هو (0.996) ، وهنا نلاحظ من هذا المثال أن هذا الإرتباط العالي بين المتغيرات التوضيحية يمكن ان يجعل معلمات الإنحدار ليست ذات معنوية إحصائية عالية أو يعكـــس إشـــارات المعلمات على نحو مخالف لمنطق النظرية الإقتصادية ، إن مشـــكلة الإرتبـــاط الخطي المتعدد غير مقتصرة على متغيرين توضيحيين وإنما يمكن لهذه المشكلة أن تكون موجودة بين متغيرات توضيحية عدة .

وفي هذه الحالة اذا تم القيام بإنحدار (X_1) على المتغيـــر (X_2) ســـوف نحصل على :

$$X_1 = 4.191 + 0.134X_2$$
$$t \rightarrow \quad (8.74) \quad (88.11)$$

إن قيمة إحصاءة (t) ذات معنوية إحصائية عالية جداً وإن قيمة معامـــل التحديد المعدل هي (0.993) كلها تؤشر علاقة قريبة من التامة بين المتغيرين ، فإذا عوضنا عن هذه العلاقة بين المتغير في معادلة النمـــوذج الثالـــث فســـوف نحصل على :

$$\hat{E} = 7.29 + 27.58X_2 - 151.15(4.191 + 0.134X_2)$$
$$= -626.18 + 7.33X_2$$

وهذه النتيجة قريبة جداً من قيم النموذج الأول الذي تم تقديره .

مراجع الكتاب

مراجع الكتاب

1- Christ , C, Econometric Models and Methods , John Wiley and Sons , Inc , New York , 1966 .

2- Chow , C.C. , Econometrics , McGraw-Hill , New York , 1983 .

3- Draper , N.R and H.Smith , Applied Regression Analysis , John Wiely and Sons , Inc , New York , 1998 .

4- Epstein , R.J , A History of Econometrics , North-Holland , Amsterdam , 1987 .

5- Fisher , f , M , The Identification Problem , McGraw-Hill , 1966 .

6- Greene , W , Econometrics Analysis , Mcmillan , 3rd edition , New York , 1997 .

7- Goldberger , A.s , Econometric Theory , John Wiely and Sons , Inc , New York , 1964 .

8- Griliches , Z , and Intriligator , M.D (ed) Handbook of Econometrics , Vol III , Elsevier , Amsterdam , 1986 .

9- Goldberger , A.s , Introductory Econometrics , Harvard University Press , Cambridge , 1998 .

10- Gujarati , D.N , Basic Econometrics , McGraw-Hill , New York , 1995 .

11- Intriligator , M.D , Econometrics Models , Techniques and Applications , North Holland , Publishing Company , Amsterdam , 1978 .

12- Intriligator , M.D , Mathematical Optimization and Economic Theory , Englewood , cliffs , Prentice-Hall , Inc N,J. 1971 .

13- Kane , E,J. , Economic Statistics and Econometrics , Harper International , 1968 .

14- Kennedy , P. , A Guide To Econometrics , Cambridge , Mass : MIT Press , 1995 .

15- Klien , L. , R. , A Textbook of Econometrics , Row-Peterson , 1953 .

16- Kmenta,J., Elements of Econometrics, Macmillan, New York, 1971 .
17- Koutsoyiannis , A. , Theory of Econometrics , ELIBS with Macmillan , 1977 .
18- Leser , C. , Econometrics Techniques and Problems , Griffin , 1966 .
19- Maddala , G. , Introduction to Econometrics , Macmillan , New York , 1990 .
20- Mirer , T.W. , Economic Statistics and Econometrics , Englewood Cliffs , N.J. Prentice-Hall , 1995 .
21- Pindyck , R.S. , and Rubinfeld , D.L. , Econometric Models and Economic Forcasts , 1991 .
22- Plackett , R.L. , Regression Analysis , Clarendon Press , Oxford , UK , 1960 .
23- Rao , C. , R. , Linear Statistical Inference and its Application , 2nd ed , Wiley , New York , 1960 .
24- Rao P. , and Miller , R.L. , Applied Econometrics , Wadsworth Publishing Co , Inc. Belmont , Calif , 1971 .
25- Schmidt , P. , Econometrics , Mareel-Dekkar , New York , 1976 .
26- Schneider , H. , and Barker , G.P. , Matrics and Linear Algebra , 2nd ed , Reinhart and Winston .
27- Stewart , J. Understanding Econometrics , 2nd ed , Hutchinson , London , 1984 .
28- Stewart , J. Econometrics , Philip Allan , New York , 1991 .
29- Suits , D.B. , The Theory and Application of Econometric Models , Athens , 1963 .
30- Thiel , H. , Principles of Econometrics , North-Holland , 1972 .
31- Weisberg , S. , Applied Linear Regression , Wiley , New York , 1985 .

32- Wonnacott , R.J. , and Wonnacott , T.H. , Econometrics , Wiley , 1970 .

33- Wooldridge , J.M , Econometric Analysis of Cross Section and Panel Data , The MIT Press , Massachusetts , 2002 .

34- Wooldridge , J.M , Introductory Econometrics : A Mpdern Approach , Cincinnati , Ohio , South-Western , 2000 .